nominalizações

[-D/-R] complementos sentenciais:

I know that he comes.

Roberto Gomes Camacho

Classes de palavras na perspectiva da Gramática Discursivo-Funcional

O papel da nominalização no *continuum* categorial

Classes de palavras na perspectiva da Gramática Discursivo--Funcional

ROBERTO GOMES CAMACHO

Classes de palavras na perspectiva da Gramática Discursivo-Funcional

O papel da nominalização no *continuum* categorial

Agradecimentos

Em primeiro lugar, agradeço profundamente aos membros da banca examinadora da tese de livre-docência de que resulta este trabalho, professores Ataliba Teixeira de Castilho (Unicamp), Maria Helena Moura Neves (Unesp), Francisco da Silva Borba (Unesp), Maria Luiza Braga (UFRJ) e Maria Stella Bortoni-Ricardo (UNB), cujas críticas e sugestões permitiram-me aperfeiçoar tanto a forma quanto o conteúdo do texto final.

Agradeço também à Erô pelo apoio afetivo, e a ela e à colega Cristina Carneiro Rodrigues, pela leitura atenta que fizeram da versão final do texto, especialmente à Cristina, pelas inúmeras sugestões de normativização. Ocioso reconhecer que os problemas que persistem são de minha inteira responsabilidade.

Ao professor Kees Hengeveld, da Universidade de Amsterdã, supervisor de meu estágio de pós-doutorado, pela disponibilidade em discutir pontos específicos da versão preliminar deste trabalho.

À professora Evelien Keizer, editora dos *Webpapers in Functional Grammar,* da Universidade de Amsterdã, que, incansavelmente, debateu problemas do artigo de que resultou uma parte substancial deste texto.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro concedido, na forma de bolsa

de Produtividade em Pesquisa, ao projeto *A estrutura argumental dos nomes deverbais*.

À Fundação de Auxílio à Pesquisa do Estado de São Paulo (Fapesp), pela bolsa de Pesquisa no Exterior, concedida em 2005, para o aperfeiçoamento do projeto *A estrutura argumental dos nomes deverbais*, que serviu de base para a consecução deste trabalho.

Sumário

Prefácio

Roberto Gomes Camacho atinge com este trabalho sua plena maturidade intelectual, duramente conquistada ao longo de uma carreira exemplar.

Licenciado em Letras pela Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp) em 1973, ele obteve seu mestrado em Linguística pela Universidade Estadual de Campinas (Unicamp), em 1978, e seu doutorado em Linguística e Língua Portuguesa pela Unesp, em 1984.

Tendo assumido em 1978 o posto de professor em Linguística em sua universidade-mater, ele passou a liderar seus colegas na direção da Gramática Discursivo-Funcional, implantando no programa de pós-graduação linhas de pesquisa em Descrição funcional de língua oral e escrita e em Variação e mudança linguística. Em 2005, desenvolveu um programa de pós-doutorado na Universidade de Amsterdã, tendo obtido em 2009 o título de livre-docência com o texto que se vai ler.

Atualmente é professor-adjunto da Universidade Estadual Paulista "Júlio de Mesquita Filho", onde atua na Graduação e na Pós-Graduação, desenvolvendo pesquisas nos seguintes temas: articulação de orações, classes de palavras, estrutura argumental e tipologia. É bolsista de Produtividade em pesquisa do Conselho

Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e coordenador da área de Linguística da Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp).

Importante recordar os encontros científicos que, juntamente com seus colegas funcionalistas, passou a organizar em sua universidade, entre os quais se destacam, em 2003, o International Course on Role and Reference Grammar e, em 2006, o 3rd International Course on Functional Grammar, além do grupo de trabalho sobre Marcadores discursivos.

O assunto deste livro vem sendo investigado por Roberto Camacho no interior do Projeto de Gramática do Português Falado, de que resultou, em 2008, o capítulo "O substantivo", escrito em colaboração com Marize Hattnher e Sebastião Carlos Leite, publicado por Rodolfo Ilari e Maria Helena Moura Neves na versão consolidada, Gramática do português culto falado no Brasil (vol. 2, 2008, p.21-80).

Roberto Camacho enfrenta neste livro um problema clássico do Funcionalismo: a questão da categorização linguística, focalizada no modelo em sua natureza contínua, não opositiva, não negativa. Surgiu daqui a teoria dos protótipos, já bem conhecida no Brasil.

É precisamente o lado complexo, no sentido de "interligado", que o autor considera neste trabalho. Para isso, ele toma como parâmetro as diferenças entre o nome e o verbo, situando a meio caminho o importante processo da nominalização, "postulada como constituindo um termo complexo entre os dois outros opostos, tomados como membros prototípicos". Para comprovar essa postulação, o autor argumenta que as expressões nominalizadas partilham propriedades nominais e verbais, comprovadas, neste caso, pela manutenção da estrutura argumental encontrada no verbo.

O processo da nominalização é, a seguir, cuidadosamente examinado no interior da Teoria da Gramática Funcional de Dik e da Gramática Discursivo-Funcional (GDF) de Hengeveld, Mackenzie e outros. Vindo aos dados, ele observou a predominância de expressões nominalizadas que projetam dois argumentos, se comparadas àquelas que projetam um só argumento, expresso ou não nos

enunciados. Conclui-se que as expressões nominalizadas preservam o número de argumentos do verbo de que derivam. Segue-se uma generalização sobre os achados, nos termos da teoria adotada.

Como bom cientista, Camacho reconhece que "A verdade é que a ciência não é uma magia e estar certo ou equivocado é uma questão relativa como, aliás, é verdadeiro também para a categorização linguística". Indo por aqui, não há dúvida de que seu texto estimula o debate entre as teorias linguísticas ora em curso no Brasil.

A Gramática Discursivo-Funcional surgiu justamente num momento em que estão sendo desafiadas antigas convicções dos linguistas, entre elas as que postulam as línguas naturais como um conjunto de sistemas para cuja descrição e interpretação devemos admitir um sistema central, tratando os demais como derivações, num andamento unilinear. Para sustentar isso, é preciso aceitar que o sistema central independe dos outros.

Metidas nessa camisa de força, as diferentes teorias oscilam entre candidatar a Gramática a esse posto de honra, por sua antiguidade no palco, em que outros instalam ora a Semântica, ora a Pragmática, ora o Léxico. A Gramática Discursivo-Funcional não difere de outras teorias neste particular. Ela também postula um "modelo hierárquico de gramática, constituído por um sistema multifuncional de níveis e camadas". Aparentemente, retornamos assim aos estruturalistas, que tinham introduzido a questão dos níveis na teoria linguística, pois postulavam a língua como um código. Depois, quando já se tinha substituído a língua-código pela língua-uso, houve uma explosão descontrolada na postulação de níveis, cuja aceitação leva à escolha de um nível central e à neutralização de um nível por outro. Com seu efeito que nada tem de secundário: a exclusão de um nível por outro, a negativa em reconhecer que tudo na língua funciona simultaneamente, não linearmente.

Entretanto, seria injusto considerar que a organização em camadas e níveis hierárquicos proposta pela GDF atribuiria um caráter demasiadamente estático aos procedimentos descritivos e interpretativos. E, com efeito, ao mencionar a natureza psicológica da implementação dinâmica da gramática no tocante à nominalização,

o autor deixa suficientemente claro que, também para a GDF, tudo na língua funciona simultaneamente. Deixa claro também ser ilusório pensar que os níveis superiores de organização, que regem a Formulação – o Interpessoal, que abriga motivações pragmáticas, e o Representacional, que abriga motivações semânticas –, sempre modelariam a forma final dos níveis inferiores, o Morfossintático e o Fonológico, que regem a Codificação. A preocupação do modelo teórico com o princípio de adequação tipológica, já defendido por Dik, mostra que há processos morfossintáticos que se explicam exclusivamente com base em outros processos igualmente morfossintáticos. Além de mostrar a independência dos níveis, esse princípio mostra claramente que, se a teoria é estática, ela o é apenas como a câmara escura de uma máquina fotográfica, que, por assim dizer, "congela" a dinamicidade sempre presente na atividade humana. Assim o faz a GDF com o processo sempre dinâmico que caracteriza a linguagem, mas "congelar" numa fotografia descritiva as relações de alinhamento entre os níveis de organização não implica considerar um modo estático de ser para a linguagem humana.

Essa coincidência de pontos de vista mostra que uma mesma forma de produzir conhecimento científico cimenta os alicerces de teorias aparentemente tão diversas.

Ora, justamente o princípio da categoricidade elaborado por Roberto Camacho postula um *continuum* integrativo, conduzindo o raciocínio para outra direção, toda feita de inclusões, de simultaneidades. Um dos principais méritos deste livro será desencadear esse debate, para além de fazer-nos entender melhor o processo da nominalização.

Ataliba Teixeira de Castilho

Palavras iniciais

Na tradição gramatical, a divisão do discurso em partes chamadas classes de palavras baseia-se em critérios nocionais de natureza discreta que acabam por fornecer uma abrangência supostamente universal, como se categorias próprias das línguas clássicas indo-europeias pudessem aplicar-se a todas as línguas naturais.

Todavia, a realidade é que a conhecida diversidade tipológica das línguas, com base na relação raramente biunívoca entre as categorias ontológicas e as categorias gramaticais disponíveis, não tem deixado livre de problemas nem mesmo a suposta distinção universal entre nome e verbo. Com efeito, a literatura funcionalista sobre tipologia está repleta de evidências que comprovam essa afirmação. Num trabalho bastante ilustrativo, Hengeveld (1992) postula que os sistemas de classes de palavras podem ser ordenados na escala contida em (0-1), conforme as línguas disponham ou não das quatro classes de palavras predicadoras.

(0-1) Verbo > Nome > Adjetivo > Advérbio

(Hengeveld, 1992, p.70)

Essa hierarquia pressupõe que uma categoria de predicados é mais provável de ocorrer como uma classe de palavras separada se ela estiver localizada mais à esquerda do *continuum*; pressupõe,

além disso, que cada estágio do *continuum* pode servir de ponto de corte para os diferentes tipos de línguas.

O português, assim como a maioria das línguas ocidentais, dispõe de uma classe separada de palavra para cada ponto do *continuum* em (0-1). Há, todavia, línguas, como o mandarim, que dispõem apenas de nomes e verbos, e outras, como o tuscarora, língua ameríndia da América do Norte, que dispõem somente de verbos.

É por isso que os pesquisadores que reconhecem alguma universalidade na distinção categorial entre nomes e verbos apoiam-se não numa categorização em unidades discretas, nitidamente distintas, mas numa categorização prototípica, que prevê uma continuidade categorial. Em termos tipológicos, é possível predizer que certas percepções prototípicas de entidades próximas a coisas são codificadas numa forma gramatical identificável como nomes, ao passo que percepções prototípicas de ações ou eventos são gramaticalmente codificadas como verbos (cf. Hopper; Thompson, 1984).

As classes de nomes e de verbos dispõem, assim, de correlatos semânticos que correspondem aproximadamente a entidades cognitivamente percebidas no mundo real. Para os nomes, a entidade é qualquer coisa similar a objeto, ou a uma percepção do que Givón (1979) designa "estabilidade temporal" (*time-stability*). Ao contrário, os verbos são prototipicamente ações ou eventos: representam percepções sem estabilidade temporal.

Essa dualidade se deriva de uma tendência universal de associar entidades temporalmente estáveis com a classe gramatical dos nomes e entidades não temporalmente estáveis com a classe gramatical dos verbos. Como essa correlação tem validade universal, Hopper e Thompson aplicam à noção de categorialidade o princípio cognitivo da prototipicidade, desenvolvido por Rosch (1973), segundo o qual a categorização humana não é arbitrária, mas procede de exemplares mais centrais para exemplares mais periféricos de categorias, sendo prototípicos justamente os exemplares centrais que parecem mais salientes aos falantes.

Um aspecto teoricamente instigante, relacionado às classes de palavras, é a caracterização categorial das nominalizações, que, a rigor, não podem ser consideradas membros prototípicos nem da

classe dos nomes nem da classe dos verbos. Consistem, na realidade, em categorias complexas, intermediárias num *continuum* funcional com os polos ocupados pelo nome e pelo verbo. Como nomes, deveriam referir-se a entidades perceptíveis do mundo, mas, como os verbos, podem ser providos de valência e representar não entidades de primeira ordem, que são percepções cognitivas de objetos físicos, palpáveis, mas entidades de segunda ordem, ou estados de coisas. Nesse aspecto, uma boa hipótese a defender é que quanto mais uma nominalização preservar a estrutura argumental herdada do verbo *input*, tanto mais próxima estará da referência a um estado de coisas.

De acordo com a definição de classes de palavras, proposta por Hengeveld (1992), o português pode ser classificado como uma língua diferenciada, já que dispõe de classes gramaticais distintas para o desempenho de diferentes funções. Por disporem de certa rigidez formal, línguas diferenciadas precisam contar com operações transcategoriais para derivar nomes de verbos, o que costuma ocorrer, por exemplo, no português, com processos produtivos de sufixação.[1]

Na visão mais tradicional, a nominalização envolve tanto a aquisição de propriedades nominais como a perda de propriedades verbais, numa relação de oposição discreta. Entretanto, segundo Malchukov (2004), a pesquisa tipológica mais recente tem reconhecido que a articulação entre os dois processos acarreta possibilidades mais graduais que discretas e, portanto, empiricamente mais adequadas. E, com efeito, as operações transcategoriais, quando aplicadas à nominalização, envolvem tanto descategorização quanto recategorização.

O termo *descategorização*, introduzido inicialmente por Hopper e Thompson (1984), indica que verbos e nomes perdem algumas das propriedades morfossintáticas associadas com suas funções primárias de relatar eventos e referir-se a termos, respectivamente. Assim, um verbo usado como expressão referencial adquire propriedades nominais, como caso, determinação, modificação etc., e

1 Segundo Basílio (1989), os nomes abstratos, formados a partir de verbos, servem para atender a exigências sintático-semânticas do discurso, o que justifica a enorme produtividade dos processos de nominalização a partir de verbos.

é por isso que o termo *nominalização* reúne propriedades relacionadas tanto ao processo de deverbalização quanto ao de nominalização (cf. Malchukov, 2004).[2]

Em seu funcionamento, os processos de recategorização (R) e descategorização (D) preservam uma independência relativa: como demonstram Hopper e Thompson (1984), do mesmo modo que descategorização não acarreta necessariamente recategorização, também recategorização não acarreta necessariamente descategorização. Usando exemplos do inglês, Malchukov menciona quatro possibilidades na tipologia das construções completivas encaixadas, contidas em (0-2):

(0-2)	a	[-D/-R] complementos sentenciais:	*I know that he comes*.
	b	[+D/-R] infinitivos:	*I want him to come early.*
	c	[-D/+R] nominalização oracional:	*I disapprove of his driving the car so carelessly*
	d	[+D/+R] nominalizações:	*I didn't see John's arrival.*

(Malchukov, 2004, p.57)

A construção (0-2a) se refere a um complemento sentencial: embora ocorra numa posição de objeto (SN), ela retém a estrutura de uma sentença e raramente adquire traços nominais.

No caso (0-2b), os infinitivos retêm a sintaxe interna de um SV em relação a alguns traços semânticos, como possibilidade de argumento, aspecto e voz, mas não podem assumir outras propriedades tipicamente verbais, como tempo e modo, e, no caso citado do inglês, concordância e combinação com um sujeito nominativo.

O caso (0-2c) combina, por seu lado, traços nominais e verbais: é capaz de receber objeto, enquanto o sujeito é expresso como um SP, isto é, na forma gramatical típica de possuidor.

---

2 Dado que o termo "nominalização" designa um processo de transposição categorial de que resultam outros nomes como os deadjetivais, reserva-se aqui o uso da expressão "nome deverbal" para designar o resultado do processo de transposição a partir de um predicado verbal.

Finalmente, em (0-2d), o verbo é completamente assimilado a um nome: assume todas as flexões nominais e raramente preserva traços gramaticais de verbo. Enquanto complementos sentenciais e nominalizações, representando os polos opostos na escala, são codificados por classes lexicais mais ou menos universais, como verbo e nome respectivamente, o estatuto categorial das classes intermediárias difere muito de uma língua para outra.

Vale acrescentar que os exemplos do inglês servem apenas para ilustrar as construções encaixadas possíveis e, por isso, não podem ser tomados como universais, principalmente em função da diversidade tipológica das línguas; na realidade, o que constitui um traço universal, que os exemplos ilustram bem, é o caráter contínuo, não discreto das categorias. Assim, a gramática do português, por exemplo, não licencia a ocorrência da categoria intermediária (0-2c), e só permite construções não finitas com predicados volitivos (0-2b), se houver identidade de sujeito entre matriz e encaixada, como em (0-3).

(0-3) *Eu quero ir para casa cedo hoje.*

Nesse aspecto, difere da gramática do inglês, que, como se vê em (0-2b), licencia a forma não finita mesmo com identidade entre objeto da matriz e sujeito manifesto como anáfora zero na encaixada.

Segundo Dik (1997), uma nominalização é uma construção encaixada que tem propriedades em comum com um termo nominal primário, como *demissão* em (0-4).

(0-4) *Maria lamentou a demissão de Pedro.*

Com efeito, em (0-4), o constituinte na função de paciente da nominalização aparece sob forma de sintagma de possuidor,[3] que

---

3 Conforme será tratado mais adiante, forma de sintagma de possuidor nem sempre significa função semântica de possuidor. Nomes denotando membros prototípicos contêm, de fato, a noção semântica de posse, como em *o livro de Maria*. Nesse caso, nominalizar um predicado verbal é aproximá-lo da estrutura de um nome, ajustando os termos do predicado à estrutura de um

é especialmente adequada para expressar relações no interior de sintagmas nominais. O fato de *Pedro* aparecer em (0-4) sob a forma de sintagma de possuidor é interpretado por Dik como uma propriedade nominal, e o SN *a demissão de Pedro* é descrito como uma construção encaixada na posição de complemento da oração matriz que tem como núcleo o predicado *lamentou* (cf. Dik, 1997, p.57-8).

A principal implicação teórica dessa posição é que nominalizações dispõem de valência potencial como outras construções encaixadas. Como um tipo derivado de nome, os deverbais podem ser mono, bi e trivalentes, na mesma medida em que são os verbos que lhe dão origem, conforme entende Dik (1985; 1997). Entretanto, essa interpretação de valência por herança natural não é destituída de controvérsia justamente porque a organização sintática de um nome deverbal raramente expressa todos os constituintes de sua estrutura argumental, como se observa em (0-5a).

(0-5) a *se eu (fizer) este gato e deixasse durante doze mil anos... ele vai continuar sendo um gato sem valor... não tem:: nenhuma... um valor artístico esta* *representação* (EF-SP-405)

b $(e_i\ [(f_i\text{: representação}_N\ (f_i))\ (x_{i}\text{: eu}\ (x_i))_A\ (x_j\text{: gato}\ (x_j))_U]\ (e_i))$

O nome grifado em (0-5a), cujo esquema de predicação (cf. Dik, 1989) pode ser representado em (0-5b),[4] constitui um predicado cujos argumentos potenciais são desprovidos de codificação fonológica.

Essa característica da nominalização subsidia a hipótese postulada por Mackenzie (1985; 1996) de que a predicação nominalizada

membro prototípico da classe. Esse ajuste implica que as relações argumentais sejam representadas por um SP introduzido por *de*. É esse o significado da expressão "sintagma de possuidor".

4 A representação formal, conforme se verá adiante, segue regras recursivas próprias da notação tradicional da lógica dos predicados do tipo seguinte para predicados de um lugar: $(x_i\text{: f}\ (x_i))$, em que se lê 'um x_i tal que x_i é f', e de predicados de dois lugares uma expressão como $(x_i\text{: f}\ (x_i)\ (x_j))$, em que se lê 'um x_i tal que x_i tem uma relação f com x_j' (Dik, 1989, p.63); U subscrito representa Undergoer (paciente).

é, na realidade, avalente, posição sustentada também por Basílio (1989), que alega ser a estrutura argumental obrigatória nos verbos e opcional nos nomes.

Tem-se, assim, um impasse: que hipótese é capaz de explicar mais fielmente os dados empíricos, a hipótese de preservação de valência de Dik ou a hipótese de redução de valência de Mackenzie? Neste trabalho, sustento a hipótese de preservação de valência, com base no postulado de que as nominalizações se referem não a entidades de primeira ordem, mas a entidades superiores. Como nomes deverbais que são, representam estados de coisas, preservando, portanto, a estrutura argumental do predicado verbal correspondente, conforme sugestão de Dik (1985), aplicada também por Neves (1996) e Camacho e Santana (2004).

A defesa desse postulado, com base em evidências extraídas do português falado, já foi sustentada por Santana (2005); como se verá, entretanto, pretendo impor uma interpretação alternativa para a formalização que desaguará, por conseguinte, numa interpretação teórica diferente da desenvolvida por Santana (2005).

Há razões funcionais para considerar que, na ocorrência contida em (0-5a), os argumentos, expressos por anáfora zero na estrutura interna do SN nucleado por *representação*, retomam a referência dos termos *eu* e *gato*, para o primeiro e o segundo argumento, respectivamente, um fenômeno pragmático de identificação de referentes, amplamente conhecido por caracterizar também a estrutura valencial de predicados verbais. Conforme diz Keizer (2004b, p.17):

> É essa função de identificação de referente que explica por que podem ser omitidos os argumentos de predicados nominais básicos ou derivados, ao passo que, no nível da predicação, eles devem ser sempre virtualmente expressos.[5]

5 Citações literais com tradução minha terão sempre o texto original em nota de rodapé. *"It is this function of referent identification which explains why it is that the arguments of basic or derived nominal predicates can be omitted, while at predication level they must virtually always be expressed."*

Concordo com a posição de que as expressões identificadoras no nível do termo são argumentos não expressos por razões pragmáticas, já que, de um ponto de vista lexical, mantêm o mesmo grau de obrigatoriedade do domínio verbal; mas, como veremos na análise dos dados, há também razões semânticas para não expressá-las.

Em face dessa aproximação potencialmente relevante entre classes de palavras e processos de (des-/re-)categorização, o objetivo deste trabalho é mostrar que as classes de palavras são pontos em um *continuum* categorial. Para tanto, este trabalho se limita ao estudo das duas classes de palavras, nome e verbo. Por atuarem como constituintes indispensáveis aos atos locutórios primordiais da produção da sentença, que são, respectivamente, referir e predicar, nomes e verbos são, em termos tipológicos, candidatos potenciais a universais linguísticos.

A continuidade categorial, como um universal linguístico, com diferentes pontos de corte para as diferentes línguas, pode aplicar-se também, no interior de uma mesma língua, a um termo complexo servindo de elo entre os extremos, em que se situam os membros mais prototípicos. Nas regiões intermediárias, situam-se os membros menos prototípicos e categorialmente complexos, estatuto que defendo como traço constitutivo das nominalizações extensíveis a formas não finitas do verbo, como infinitivos, gerúndios e particípios.

Como a nominalização tem uma continuidade funcional entre referir e predicar, essa continuidade depende crucialmente de assumir a hipótese secundária de que a estrutura argumental é parte constitutiva dela, quando se referindo a entidades de ordem superior. Consequentemente, quanto maior o grau de ajuste de um nome deverbal à expressão formal e semântica de um termo de primeira ordem, tanto mais próximo é seu significado de "entidade" e, portanto, mais próximo da "nominidade" prototípica. Também a recíproca é verdadeira: quanto mais ele preserva a estrutura argumental do verbo original, tanto mais próximo está da "verbidade" prototípica. Postulo que a preservação da valência formal na nominalização implica, correlativamente, manter o estatuto semântico derivado da grade de relações temáticas herdada do verbo *input*.

A dotação ou não de estrutura argumental é um aspecto crucial da argumentação em favor da hipótese que pretendo defender; por isso, a investigação da forma de expressão da estrutura valencial das nominalizações deverá permitir identificar também os princípios que motivam a escolha de um ou de outro termo e, portanto, de que modo se resolve a competição entre os argumentos para ocupar a posição de sintagma de possuidor prevista por Dik (1985).

As nominalizações são, portanto, construções potencialmente valenciais, na mesma medida em que o são outros tipos de construções encaixadas, como infinitivos e particípios. Embora essa posição seja compartilhada por Dik, ao tratar de construções encaixadas, ele não faz qualquer referência à natureza aberta ou fechada das nominalizações atuando como núcleo de predicações, que acaba sendo uma propriedade relevante para explicar a natureza valencial delas.

A nominalização constitui um dos mais estudados tipos de operação transcategorial, como se vê nos trabalhos de Comrie e Thompson (1985), Noonan (1985), Lehmann (1988), Dik (1985; 1997), Mackenzie (1985; 1996), Croft (1991), Koptjevskaja-Tamm (1993), Cristofaro (2003) e Malchukov (2004).

O interesse principal desses estudos é dar um tratamento funcional ao enfoque tipológico; nesse âmbito, principalmente os três últimos dialogam bastante entre si, embora o objeto de estudos de um não recubra necessariamente o do outro. Koptjevskaja-Tamm (1993) faz um estudo tipológico dos nominais de ação na função de complementação, enquanto Cristofaro (2003) enfoca, além da complementação, também a subordinação adverbial e a relativa. Malchukov (2004), por exemplo, argumenta que seu trabalho, além de fornecer uma explicação coerente e unificada para as nominalizações de ação como mecanismos de operação transcategorial, com base numa amostragem suficientemente ampla, consistiria, ao mesmo tempo, numa ponte entre as abordagens funcionalista e formalista.

Os trabalhos que tratam da estrutura argumental de nominalizações, principalmente no interior do arcabouço da Teoria da Gra-

mática Funcional (doravante GF), desenvolvido por Dik (1989; 1997) ou de outros enfoques funcionalistas, preocupam-se mais com questões relacionadas ao componente semântico e ao gramatical do que com questões relacionadas ao componente pragmático. Este deveria, justamente, contar como o domínio mais abrangente de explicação, já que um enfoque funcional dá ênfase especial ao caráter não autônomo da morfossintaxe na explanação teórica dos fenômenos linguísticos.

Nesse caso, os níveis pragmático, semântico e morfossintático do componente gramatical acabam tendo relações de autonomia por contarem com categorias próprias, mas também relações de interdependência por conta de codificação morfossintática motivada por fenômenos derivados do nível pragmático e do semântico. Além disso, pouco trabalho se devota para a análise de dados reais de conversação, uma das razões de ser do enfoque funcionalista.

De um ponto de vista metodológico, este trabalho está seriamente comprometido com motivações derivadas do Nível Interpessoal e do Representacional para a codificação no Nível Morfossintático, de acordo com o modelo de Gramática Discursivo-Funcional (doravante GDF), que foi praticamente instaurado por Hengeveld (2004a) e consolidado por Hengeveld e Mackenzie (2008).

Esse compromisso o insere definitivamente no âmbito dos trabalhos funcionalistas, além disso, de um ponto de vista empírico, está seriamente comprometido com a interpretação de dados reais de conversação, o que também é, ou pelo menos deveria ser, outro compromisso com a metodologia funcionalista. Esse tratamento empírico envolve uma análise quantitativa e qualitativa dos dados de base especialmente voltada para sustentar a confirmação da hipótese acima formulada de continuidade categorial.

Este trabalho está formalmente distribuído em três grandes partes, que correspondem *grosso modo* a uma investigação teórica, a uma investigação empírica e às implicações teóricas resultantes.

A Parte I está separada em três capítulos de apresentação de um ponto de vista teórico sobre o assunto abordado. Como pretendo enfocar as classes de palavras lexicais, principalmente o nome e o

verbo, defendo o princípio de que, de uma perspectiva funcionalista, as partes do discurso constituem categorias escalares, contínuas. Esse assunto está particularmente desenvolvido no Capítulo 1.

Esse capítulo se inicia com o questionamento da universalidade da divisão clássica das partes do discurso, que remonta a pelo menos cinco séculos a.C. Meu principal objetivo, nesse esboço evolutivo, é não apenas demonstrar que a universalidade das classes tradicionalmente postuladas é mais aparente que real, mas também postular que, de um ponto de vista tipológico, a distribuição das classes pelas línguas do mundo é também uma questão de gradiência. Isso significa que, embora algumas partes do discurso, como o verbo, sejam compartilhadas por praticamente todas as línguas, com direito a ser quase um universal, a distribuição transistêmica mostra que são dotadas antes de continuidade funcional que de diferença discreta.

A consequência mais evidente do postulado de gradiência categorial é que as classes são representadas por membros mais prototípicos e, com base nessa posição teórica, pretendo tratar o português tipologicamente como uma língua diferenciada, um tipo intermediário entre uma classe rígida de línguas e uma classe flexível.

As evidências empíricas para sustentar essa noção de gradiência categorial serão especialmente buscadas no âmbito de uma categoria complexa, a das nominalizações, mais especificamente a do nome deverbal, uma categoria que fica a meio caminho entre referência e predicação.

Como o principal suporte teórico é o da GF, e sua mais recente versão, a GDF, o Capítulo 2 está voltado para a explicitação desse modelo; essa tarefa consiste, basicamente, em fornecer um esboço da evolução recente dessa abordagem, tarefa que obriga necessariamente a detalhar a organização em níveis e camadas que o modelo formula para a gramática das línguas naturais.

Antes de chegar aos dados do português falado, que fornecem um mínimo de evidência empírica ao postulado acima mencionado, a continuidade do enfoque teórico se dará no Capítulo 3 pelo tratamento devotado à nominalização no arcabouço da Gramática

Discursivo-Funcional. Nesse âmbito, discuto detalhadamente as hipóteses alternativas sobre a preservação ou redução de valência no nome deverbal. O objetivo específico é distinguir as duas alternativas em função da posição teórica que assumo neste trabalho com base em evidências empíricas que comprovam a ideia de que a valência é parte integral da nominalização.

A Parte II se destina à apresentação de evidência empírica para a hipótese de preservação da estrutura argumental na nominalização. No Capítulo 4, invoco evidências empíricas, fortemente arraigadas nos dados, para defender a hipótese, crucial para esta tese, de que as nominalizações que se referem a entidades de nível superior são dotadas de estrutura argumental, que pode ou não ser especificada formalmente em função de determinações semânticas e pragmáticas. Forneço, ainda, os fatores que movem a competição entre os argumentos e o modo como os argumentos são expressos no uso.

A Parte III trata das generalizações e implicações teóricas. Fazendo um elo com a investigação empírica, no Capítulo 5, procuro fornecer uma interpretação das construções nominalizadas na Gramática Discursivo-Funcional (GDF), tentando atar com a interpretação funcional e formal alguns fios que possam ter ficado soltos à espera de amarração na Parte I. Forneço um esboço de formalização que mostra ser a nominalização um caso exemplar do princípio de adequação psicológica da GDF. Para tanto, baseio-me no princípio de implementação dinâmica e de organização em níveis e camadas, que acabam por fornecer uma explicação teórica altamente plausível para a ausência de especificação formal de valência da nominalização.

O Capítulo 6 fecha a Parte III com a apresentação das generalizações e implicações teóricas propriamente ditas, decorrentes dos resultados da análise e, ao mesmo tempo, da proposta de interpretação da nominalização no arcabouço da Gramática Discursivo-Funcional. Nesse capítulo, discuto especificamente o papel que a nominalização exerce para a continuidade categorial das classes de palavras em virtude de seu caráter complexo, a meio caminho entre a referência a entidades e a atribuição de propriedades.

No Capítulo 6, discuto também os principais traços definidores de quatro classes de palavras predicadoras, tendo por objetivo defender a ideia de que a organização da predicação em termos do predicado verbal nuclear e do modificador adverbial dispõe de uma estrutura subjacente muito similar à do SN, em termos do núcleo nominal e da modificação adjetival. Essa organização é uma forma de atender a duas principais funções dos dois atos locutórios primordiais de linguagem, o de referência e o de predicação, que são mais bem traduzidas pela codificação em volta de nome e adjetivo, por um lado, e verbo e advérbio, por outro.

Tenho a expectativa de que esse capítulo de fechamento conecte outras pontas soltas dos fios desenrolados na Parte I sobre o caráter contínuo das classes de palavras e, por isso, também prototípico de alguns de seus membros, atributo cuja principal evidência é a nominalização representando entidades de ordem superior.

Encerrada a descrição empírica com as consequentes implicações teóricas, nas Considerações Finais, faço um balanço das principais hipóteses em relação aos resultados, e alinhavo algumas consequências relevantes das posições assumidas ao longo do trabalho. Esse fechamento constitui, portanto, uma avaliação necessária cujo compromisso principal é um acerto de contas rigoroso com a proposta de trabalho lançada nesta Introdução.

Parte I

As classes de palavras na perspectiva funcional

1
A NATUREZA CONTÍNUA DAS CLASSES DE PALAVRAS

Introdução

Os estudos tipológicos exaustivamente desenvolvidos sob a égide da linguística funcional sobre categorias lexicais mostram que os universais se restringem a hierarquias implicacionais, que, por definição, produzem categorias escalares. Esse princípio reforça a ideia de que as classes de palavras constituem categorias gradientes, o que relativiza o postulado da tradição gramatical greco-latina de que as classes são entidades discretas e universais.

Afirma Neves (2005, p.13) que "o exame da gramática ocidental tradicional necessariamente remonta à gramática grega. E é analisando a marcha da instituição da gramática grega que vamos encontrar a razão para aquelas características, que então se legitimam e se explicam". Como é nela que se pode rastrear a universalidade das classes, especialmente na filosofia estoica, pelo menos dois séculos a.C., discutirei ligeiramente na Seção 1.1 a questão da universalidade dessa divisão que, como se verá, é mais aparente que real, e procurarei enquadrar o português numa classe tipológica de acordo com a divisão em classes de palavras. Encerrarei a discussão dessa suposta universalidade fazendo, na Seção 1.2, uma pequena retrospectiva histórica sobre o caráter nocional das classes

de palavras. Em seguida, na Seção 1.3, procurarei mostrar como a questão das classes pode ser encarada de uma perspectiva funcional e tipológica, de que sobressai um tratamento contínuo para a categorização.

A enumeração das classes e os critérios para a determinação delas acompanham geralmente o paradigma que prevalece em cada período da história do pensamento linguístico. É difícil afirmar com certeza que ponto de vista teórico prevalece hoje em dia nos estudos linguísticos, principalmente no que tange à explicação da organização morfossintática da linguagem; mesmo que pareça tratar-se mais de convivência que de competição entre enfoques de natureza formal e de natureza funcional, é necessário promover-se logo a escolha de um ponto de vista. E a escolha que norteia a presente discussão sobre o lugar das nominalizações na organização lexical das classes de palavras é a de um enfoque funcionalista que, como tal, abriga princípios de natureza cognitiva, como o de prototipicidade.

Como se trata de discutir a questão da categorialidade, este capítulo não se restringe, todavia, apenas às nominalizações, mas às principais classes lexicais como nome, verbo, adjetivo e advérbio, que exercem funções de predicados na sentença; ficam, portanto, descartadas as partes do discurso que cumprem funções gramaticais, as chamadas classes fechadas como as categorias dêiticas dos pronomes e advérbios, e as categorias dos conectores em geral, e dos determinantes (artigos e pronomes demonstrativos). Apoio-me em três razões relacionadas e igualmente pertinentes para esse recorte:

(i) As classes de predicados são semanticamente plenas ou autossuficientes e, portanto, constituem formas livres.
(ii) Admite-se que haja no léxico, ao menos nas teorias da gramaticalização, em virtude do princípio de unidirecionalidade, um *continuum* que respeita a seguinte correlação: Categoria Maior [nome, verbo, pronome] > Categoria Média [Adjetivo, Advérbio] > Categoria Menor [Preposição, Conjunção] (cf. Hopper; Traugott, 1993).

(iii) Considerando-se que o princípio da unidirecionalidade esteja ao menos parcialmente correto,[1] é de se esperar que o processo de gramaticalização se aplique mais diretamente à recategorização de palavras plenas em palavras gramaticais, como atestam vários casos de advérbios ou de sintagmas adverbiais que passaram a atuar ou como conectivo, caso de *embora,* ou que ainda não deixaram de ser advérbios, como *no entanto*, *entretanto, então* etc., conforme o *continuum* mencionado em (ii).[2]

1.1 A (aparente) universalidade das classes de palavras

A classificação do léxico em categorias ou classes de palavras tem sido uma preocupação constante dos linguistas. No final do século V a.C., no plano filosófico, os gregos se interessavam pela linguagem para discutir o modo como as palavras representam as coisas, se o signo provém da natureza, *phýsei*, ou é imposto por uma convenção, *thései*, um tipo de preocupação que se dirigia mais à origem do léxico.

Segundo Neves (2005, p.48-49), é o *Crátilo,* diálogo que trata da questão da justeza dos nomes, a obra de Platão mais diretamente devotada aos problemas da linguagem. Nesse diálogo, Crátilo defende a ideia de que, como a língua é o espelho do mundo, os significados que ela representa provêm da natureza. Hermógenes

1 Castilho (2003) é o principal crítico desse princípio na linguística brasileira; por razões que não cabe aqui discutir, prefere falar em multidirecionalidade.

2 É necessário lembrar que a delimitação das classes de palavras nesse *continuum* tem muito que ver com o modo como as línguas indo-europeias organizam em categorias a percepção cultural de seus falantes. Além disso, é um pouco complicado inserir os pronomes na categoria maior, já que, embora constituam formas a meio caminho entre livres e dependentes, são decididamente categorias gramaticais e não lexicais, considerando-se pelo menos o critério clássico também questionável da natureza aberta ou fechada das classes.

defende posição contrária, sustentando que, sendo o resultado de uma convenção, a língua tem natureza arbitrária.

Assumindo uma posição intermediária, Sócrates, a terceira personagem desse diálogo, ressalta tanto os pontos fortes quanto os pontos fracos dos argumentos levantados por Crátilo e Hermógenes, encaminhando a discussão a uma solução conciliatória.

Conduzindo a argumentação pelas palavras de Sócrates, pouco a pouco Platão leva o leitor a deduzir que há um elemento de verdade tanto na posição de Crátilo quanto na posição de Hermógenes. Há, de fato, muitas palavras que possuem uma correção intrínseca em consonância com a *phýsis*, mas há outras, entendidas por convenção *thései*, em que não se pode detectar nada de essencialmente natural, por terem sido malformadas na origem ou por terem sido corrompidas pela passagem do tempo (cf. Weedwood, 2002, p.26).

Como testemunham os trabalhos da linguística moderna (cf. Haiman, 1987), há um tanto de convencional e um tanto de icônico nas línguas, verdade a que o próprio Platão já conduz seu leitor: para ele, embora ainda haja vestígios de que, em sua origem, a linguagem tenha estado ligada diretamente à realidade, é de esperar que a passagem do tempo tenha como que "corrompido" a relação perfeita entre palavras e realidade, e boa parte da linguagem tenha passado, então, a ser entendida por relação convencional.

Mantida a controvérsia de o nome ser instrumento ou imagem, o *Crátilo* acentua a ideia de que, independentemente do modo como representa a realidade, o nome deve ser separado da coisa. Assim, conforme afirma Neves, "não mais cabe investigar a condição natural ou convencional dos nomes, porque não mais cabe fixar-se neles para atingir o conhecimento. A partir de Platão, quaisquer referências a naturalismo ou convencionalismo têm de ser examinadas em um contexto mais rico e mais complexo" (Neves, 2005, p.57), em que se divisa a separação entre o nome e a coisa, contexto que se encaminha para a consideração de uma função representativa da linguagem, que seria retomada pela linguística saussuriana.

Já está presente em Platão o princípio de que o enunciado completo é dividido entre duas "partes do discurso", *ónoma* e *rhema*, que vieram a assumir sentidos técnicos até hoje amplamente cor-

respondentes a sujeito e a predicado. A base dessa divisão era originalmente funcional e semântica, não formal, como se sente ser hoje, limitada aos termos que correspondem às categorias de sujeito e predicado, mais ligadas à dimensão sintática da linguagem.

Aceita a separação entre o nome e a coisa, Aristóteles, já no século III a.C., explica o processo com o fato de que os signos falados, mediados pelos signos escritos, representam impressões da alma, que não passam de aparência das coisas reais. As impressões e as coisas são as mesmas para todos os homens, mas as palavras que representam a interpretação dessas impressões são necessariamente diferentes, uma visão que se aproxima muito da concepção moderna de "forma" e "substância" preconizada inicialmente por Saussure ([1916] 1977) e depois sistematizada por Hjelmslev (1973).

Os estoicos, nos séculos III-II a.C., sofisticaram um pouco mais as relações estabelecidas por Aristóteles, acrescentando uma etapa nesse processo entre a recepção passiva das impressões do mundo e a fala, denominando-a por *conceito*, isto é, uma noção apta a ser verbalizada. Os estoicos concordam com a posição sustentada por Aristóteles de que todos os homens podem receber as mesmas impressões das coisas que percebem, mas são exatamente os conceitos formados dessas impressões que diferem e são eles que são representados na fala.

Um conceito é representado num enunciado significativo dirigido pelo pensamento racional, *lógos*, e um enunciado articulado é um *léxis*. Diferentemente de um *lógos*, um *léxis* não precisa ser provido necessariamente de significado; assim, a palavra em estado de dicionário, com seu valor intensional, é *léxis*, ao passo que, ao entrar em combinação com outras palavras para formar um enunciado significativo, passa a ser *lógos*.

Essa distinção entre *logos*, visto como uma entidade significativa, e *léxis*, a palavra enquanto forma não sintagmaticamente atualizada, é ainda hoje fundamental para o pensamento linguístico e teve forte impacto na filosofia grega pós-estoica para o processo gradual de reconhecimento das "partes do discurso". O ponto de partida deveria sempre ser o *lógos*, o enunciado visto como significativo, e se um enunciado corresponde ou não à realidade, isto é,

se é falso ou verdadeiro, seria necessário justamente verificar que parte dele prescinde dessa correspondência com a realidade.

Como se vê, foram os filósofos posteriores a Platão, como seu discípulo Aristóteles e os estoicos, que desenvolveram e refinaram o sistema de classes de palavras. Weedwood (2002) dá notícia de um fragmento de papiro do século I d.C. que preserva os parágrafos iniciais de uma gramática que define nove partes do discurso. Em todas as definições ali contidas, predominam procedimentos semânticos não formais, tanto nos critérios usados para distinguir as partes do discurso, quanto nas próprias categorias a que se chegou.

Segundo Weedwood (2002), nota-se uma preocupação especial com o aspecto semântico na *Sintaxe* de Apolônio Díscolo, que data de II d.C. A gramática grega foi definitivamente codificada pelo erudito alexandrino Dionísio da Trácia (século II a.C.), mas são os estudos de sintaxe de Apolônio Díscolo, que realmente sobreviveram no Ocidente, filtrados pelas *Institutiones grammaticae* de Prisciano, que aplicou a doutrina do grego ao latim.

Entende Câmara Jr. (1972, p.149) que as nossas gramáticas utilizam ainda hoje "um antíquissimo quadro, que nos foi legado pela tradição greco-alexandrina e que apareceu, pela primeira vez, sob forma didática no manual do Dionísio da Trácia, 120 anos antes de Cristo". De qualquer modo, embora tenha sido o livro didático do Oriente grego, essa gramática ficou desconhecida no Ocidente até sua primeira edição impressa, em 1727 (Weedwood, 2002, p.34).

Para Apolônio Díscolo e Dionísio da Trácia, a frase é definida como *autotelēs lógos,* significando "expressão que se basta", "que tem um fim em si mesmo". Essa noção, que tinha importância fundamental na filosofia grega para estabelecer a distinção entre completude e incompletude proposicional, foi empregada por Dionísio da Trácia com um valor específico, sugerindo que "a frase é uma palavra, ou conjunto de palavras, que, para a comunicação linguística, feita em dado momento, não depende de outras palavras anteriores ou seguintes.

No entanto, a tradição gramatical que se seguiu à orientação grega filtrada, por seu lado, pelo olhar romano associou a noção de

autossustentação à ideia de frase, como uma unidade que basta a si própria, isto é, "uma reunião de vocábulos com sentido completo" (Câmara Jr., 1972, p.163-164).

Como bem ensina Câmara Jr., nem "reunião de vocábulos" nem "sentido completo" podem ser expressões definidoras de uma frase, já que *Fogo!*, por exemplo, enunciada na situação concreta em que se acha o locutor, não consiste nem mesmo numa reunião de vocábulos, quanto mais numa reunião de vocábulos com sentido completo. "Não se trata assim de um sentido completo, mas de um funcionamento autônomo no intercâmbio social" (Câmara Jr., 1972, p.164)

Confirma essa ideia a posição de Castilho, segundo a qual, com a noção de autossustentação, "os gregos estavam destacando a importância das sentenças numa situação comunicativa, num texto, em que cada uma delas tem de ter uma atuação autotélica, funcionando como um conjunto de elementos ligados solidariamente, para a constituição do sentido textual" (Castilho, 1998, p.86).

A distinção entre frase e palavra não deve ser buscada, portanto, nem na extensão formal nem na extensão do conteúdo, mas exclusivamente no plano discursivo em que opera como mecanismo de interação social. Assim, uma palavra única, como o nome próprio, *Pedro*, por exemplo, deve ser vista como uma frase no enunciado exclamativo "Pedro!" (cf. Câmara Jr., 1972, p.162).[3]

1.2 A definição nocional de classe de palavra

A classificação do léxico em classes de palavras é, como vimos, uma preocupação central da tradição gramatical do Ocidente. No geral, há muitos problemas relacionados aos critérios majoritaria-

3 Como se esclarecerá no próximo capítulo, a Gramática Discursivo-Funcional, principal base teórica deste trabalho, dá total reconhecimento teórico ao caráter autotélico dessas frases, que, como atos discursivos, são um exemplo cabal de que uma teoria linguística não pode tratar apenas de sentenças completas, a não ser que não estejam comprometidas com o discurso, como ocorre geralmente com as teorias formais.

mente nocionais de classificação, cujos traços definidores, pretensamente universais e excessivamente heterogêneos, não são mais que o reflexo de traços específicos da estrutura gramatical do grego e do latim.

Uma crítica óbvia à definição nocional simples de que um verbo denota ação é a de que basta comparar uma forma como *atacar* com um nome como *ataque*, por exemplo, que denota a mesma coisa. Outra crítica contumaz é a de que essas definições nocionais que povoam a gramática tradicional consistem num verdadeiro círculo vicioso, quando aplicadas a nomes abstratos, como *eletricidade, verdade, beleza*. A única razão, segundo Lyons (1979), para dizer que essas palavras denotam "coisas", ou entidades, é que as palavras que as exprimem são nomes.

A ambiguidade na definição das classes também é motivada pela aplicação de critérios pouco esclarecedores. Segundo Lyons (1982), rótulos denominadores de classes, como *nome, verbo* etc., são empregados com a mesma ambiguidade que o termo *palavra* na expressão *classe de palavras*.

Longe de ser um problema meramente terminológico, distinguir o item lexical "menino" de suas formas de manifestação *meninos, menina* etc. se torna um procedimento metodológico necessário, que impediria, por exemplo, o reconhecimento pela tradição gramatical de que uma mesma palavra poderia pertencer a duas diferentes classes. Esse problema é afeito principalmente a línguas de morfologia menos complexa, como o inglês, em que a forma *broken* pode ser classificada, conforme a distribuição sintática, como um adjetivo ou um verbo.[4]

O que definiria a classe, nesses casos, não é somente o núcleo semântico, que, preservado no morfema lexical da palavra, forneceria critérios para afirmação equivocada de que a mesma "palavra" pode ser categorizada em diferentes classes. Numa interpretação distribucional, é a posição sintática da forma *broken* que permitiria

4 Já em português, pelo menos o particípio em tempos compostos se distingue do adjetivo por diferenças flexionais.

incluí-la ora numa classe de particípios verbais ora numa classe de adjetivos em inglês; assim, duas formas têm a mesma função sintática se, e somente se, possuem a mesma distribuição, isto é, se são intercambiáveis em todas as sentenças gramaticais de uma língua (cf. Lyons, 1982, p.111).

O círculo vicioso se rompe quando se inclui na definição nocional do nome que essa categoria denota entidades em função dos critérios de distribuição formal. Palavras como *eletricidade, verdade* e *beleza* distribuem-se na mesma classe formal de *menino, pedra, árvore* e *cachorro*, ainda que estas pertençam, ao mesmo tempo, à categoria de palavras que denota entidades físicas. As duas subclasses não são semanticamente coextensivas, mas a classe de distribuição formal permite incluir tanto as palavras que denotam entidades físicas quanto as que não denotam (Lyons, 1979). Só a partir da noção de distribuição formal é que seria possível uma caracterização adequada e não circular.

As formas que pertencem à mesma classe de distribuição variam de acordo com a aplicação de diferentes tipos de morfemas flexionais, que acabam servindo como um critério morfológico reforçador da identificação das categorias. Assim, os nomes se flexionam em caso, número e gênero, no que são acompanhados pelos determinantes e modificadores, e o verbo é conjugado de acordo com pessoa, número, tempo e modo.

A dependência distribucional das classes, que se vê em várias partes da proposta de Lyons (1979), abre caminho para a adoção de perspectiva teórica mais recente, que veio a ser conhecida como Teoria dos Protótipos. A ausência de coextensividade entre classe nocional e classe formal, ligeiramente esboçada aqui, dá margem para pensar que nomes denotadores de entidades abstratas apenas não constituem membros prototípicos de sua classe, que seriam justamente os que denotam entidades físicas e discretas.

Nessa possibilidade alternativa, em si mesma dificilmente sujeita a controvérsias, a caracterização semântica é circunscrita a atributos criteriais que permitem a formulação de categorias prototípicas. Nas duas últimas décadas, a proposta de que objetos físicos

são protótipos de nomes e a de que ações físicas são protótipos de verbos vem sendo discutida como enfoque alternativo à abordagem formal (Lyons, 1979; Givón, 1979, 1984; Hopper; Thompson, 1984; Langacker, 1987). Um traço geral desses trabalhos é a recusa em assumir a posição já tradicional na lógica formal de que o significado é redutível às condições de verdade da proposição.

Defendendo a possibilidade de definições nocionais de categorias gramaticais básicas, Langacker (1987) afirma que os argumentos contrários mais comuns são em geral superficiais e crucialmente dependentes de princípios dúbios. Sustenta a impossibilidade teórica de estabelecer uma definição semântica para categorias de nomes e verbos somente com base em critérios de inclusão categorial, determinados seja por condições suficientes e necessárias, seja por princípios lógicos baseados em condições objetivas de verdade.

Outro argumento comum contra definições nocionais é que tanto o nome quanto o verbo podem descrever o mesmo estado de coisas, como ocorre com o paralelismo entre *atacar* e *ataque*, acima mencionado. Assume-se tacitamente, numa concepção objetiva, que, como o significado da expressão é independente de concepção humana, é absolutamente determinado pela situação que ela descreve, ou seja, suas condições de verdade. Recusando essa visão, argumenta Langacker (1987) alternativamente que, como o nome e o verbo constroem o evento a partir de imagens contrastivas, representam não o mesmo estado de coisas, mas dois estados de coisas semanticamente distintos, principalmente porque o uso da nominalização implica uma espécie de "reificação" conceptual do estado de coisas.

1.3 Prototipicidade e classes de palavras no enfoque funcionalista

O apego da linguística ocidental à categorização das classes de palavras tem, como se viu, origem na tradição filosófica, ainda que prevalecesse a ideia de que as "partes do discurso" eram elementos componentes da proposição, que deslizaram, mais tarde na gramá-

tica alexandrina, para entidades formais, entendidas como classes de palavras.

A divisão de palavras em classes numa língua parece corresponder de fato a uma divisão categórica do universo linguístico, mas não de um modo que poderia constituir algum tipo de universal linguístico.

Por isso mesmo, para Sapir, a classificação tradicional dos vocábulos em partes do discurso não passa de "uma aproximação vaga e incerta do inventário cabal da experiência" (Sapir, 1971, p.119). Para ele, é apenas uma questão idiomática, própria das línguas indo-europeias, não ser possível dizer, em português, *ela vermelha*, exatamente como se diria *ela é vermelha*, isto é, usando o adjetivo como verbo.

Para representar o mesmo significado, seria necessário usar *ela se avermelha*, o que significaria empregar um verbo deadjetival apropriado para executar a tarefa da parte do discurso chamada "verbo" e reservar para a função predicativa de *ela é vermelha* a parte do discurso chamada "qualificativo" ou "adjetivo". Todavia, argumenta Sapir, centenas de outras línguas poderiam simplesmente usar como verbo a mesma palavra que funciona como qualificativo numa construção predicativa.

Poderia haver, por outro lado, ideias transmissíveis apenas por determinada classe de palavra. Como seria possível, por exemplo, indaga Sapir, dispensar o "para" da frase "ele veio para casa"? Uma estratégia simples consiste em dizer, por exemplo, "ele alcançou a proximidade da casa" e, para "ele olhou para o espelho", seria possível dizer "ele perscrutou o interior do espelho". Está claro que essas estratégias alternativas desagradam o falante do português, porque elas não se adaptam facilmente aos mecanismos formais disponíveis pelo sistema linguístico. Há, todavia, conforme argumenta Sapir (1971, p.121), muitas línguas que expressam as relações locais mediante o uso de nomes:

> A conclusão de um tal exame seria deixar-nos convencidos que a "parte do discurso" reflete não tanto a nossa análise intuitiva da realidade quanto a nossa habilidade em compor, partindo da rea-

> lidade, uma variedade de moldes formais. Uma parte do discurso, fora das limitações da forma sintática, não passa de um fogo fátuo.
>
> Eis por que nenhum esquema lógico das partes do discurso – seu número, natureza e limite precisos – é do menor interesse para o linguista. Cada língua tem um plano seu. Tudo depende das demarcações formais que ela admite.

Sapir parece admitir que os rótulos usados pela gramática tradicional para classificar as palavras nada têm de nocional, vistos de um ponto de vista universal, mas "denotam classes gramaticais distribucionalmente justificáveis, às quais se poderiam dar quaisquer rótulos arbitrários" (Lyons, 1979, p.154).

A necessidade de tratar as partes do discurso no âmbito de classes distribucionais não universais tem a ver com a incomensurabilidade das línguas quanto à representação da percepção do mundo pelas categorias gramaticais, princípio axiomático da linguística antropológica do começo do século XX. Esse princípio não deixa imune nem mesmo a distinção entre nome e verbo, como atesta Sapir com o exemplo do paiute, língua que usa um verbo com o sujeito incorporado para denotar "lago" (1911 apud Hopper; Thompson, 1984, p.705).

Sobre o mesmo assunto, Câmara Jr. (1972) menciona o caso do waiwai, uma língua do Alto Amazonas, que manifesta a divisão dos nomes em alienáveis e inalienáveis mediante presença ou ausência do morfema *-ni-*, uma distinção da qual o português absolutamente não cogita. Câmara Jr. conclui pela inexistência de "maior relação entre a organização categórica de uma língua e as categorias lógicas que a filosofia grega procurou depreender como essenciais para o raciocínio humano" (Câmara Jr., 1972, p.120).

Nesse aspecto, a linguística antropológica norte-americana concorda plenamente com um ponto crucial da linguística estruturalista europeia, principalmente com o arcabouço teórico criado por Saussure: o de que o significado de uma forma linguística é determinado pelo próprio sistema, que fica assim imune a determinações extralinguísticas, derivadas do modo como as pessoas interagem entre si e do modo como percebem e categorizam o mundo.

A língua assim pensada é um objeto autônomo, com estrutura e princípios constitutivos próprios. Nesse caso, a noção de que a frase é uma unidade que basta a si própria percorre todo o pensamento linguístico para contaminar, inclusive, o próprio conceito de linguagem.

Essa visão arbitrária da relação entre realidade e representação linguística também deve ser relativizada em favor de algum grau de universalidade. E, com efeito, para não parecer um mero "demolidor sistemático", Sapir salva duas categorias que ele considera mais ou menos universais, o nome e o verbo.

Mesmo assim, ao afirmar que nenhuma língua prescinde totalmente da distinção entre nome e verbo, ressalva que há casos particulares em que também a natureza dessa distinção é ilusória (Sapir, 1971, p.122).[5]

Menos demolidor ainda que ele, Câmara Jr. acha possível imaginar um quadro amplo, compreensivo e coerente para usar como ponto de partida para a análise linguística as categorias e as relações já consagradas pela tradição gramatical. Dentro desse quadro, diz ele, "haverá, conforme a língua, novas delimitações; nem fica excluída a possibilidade de que falte uma das classes básicas" (Câmara Jr., 1972, p.150).

Com o advento da teoria gerativa, o princípio da autonomia, presente no estruturalismo, adquiriu um novo sentido. Embora a língua enquanto sistema de conhecimento com sede no cérebro pareça mais dependente de seus usuários, adquirindo um caráter mais mentalista, a noção de autonomia se sustenta no princípio de modularidade da mente.

Um dos módulos da mente é a faculdade de linguagem, ou competência gramatical, entendida por Chomsky (1980) como um componente computacional capaz de gerar de modo inequívoco as sentenças gramaticais da língua. Como a faculdade de linguagem é entendida com base no sentido restrito de competência gramatical,

---

5 De fato, o tuscarora, como se verá em discussão adiante, dispensa a diversidade das classes de palavras, atribuindo a uma predicação mesmo a referência a entidades (cf. Hengeveld, 1992).

que é, por sua vez, em princípio independente de outras faculdades mentais, a teoria gerativa acaba por preservar a autonomia da linguagem, consolidada no paradigma estruturalista pós-bloomfieldiano.

Entende Taylor (1989) que o princípio de autonomia em Chomsky é preservado à custa de uma extrema idealização, que extrai da linguagem os elementos que ela incorpora de outras faculdades mentais. Esses elementos são justamente a competência pragmática, o conhecimento das condições de uso conforme diferentes propósitos, e o conhecimento do sistema conceptual, que possibilita ao falante fazer operações mentais de percepção, categorização e simbolização do mundo. De um modo um tanto simplificado, pode-se dizer que a relevância semântica das classes de palavras na perspectiva gerativa é extremamente minimizada.

E, com efeito, esse modelo pressupõe um número limitado de categorias formais supostamente universais, como nome, verbo, determinante. Além disso, os itens do léxico, associados aos traços sintáticos correspondentes aos nódulos formais [N], [V] etc., são concebidos como binários, primitivos, universais e supostamente inatos.

Além disso, o significado do item lexical que preenche as posições desse marcador sintagmático abstrato é absolutamente irrelevante, já que as regras sintáticas operam independentemente da natureza semântica dos itens lexicais que preenchem as posições da categoria abstrata (Taylor, 1989, p.186-7).

Como aqui se trata das classes de palavras como um dos modos de categorização linguística, impõe-se naturalmente a questão da possibilidade ou não de lidar com sistemas de categorização dentro de concepções autônomas de linguagem, como o estruturalismo e o gerativismo. O estruturalismo acredita na arbitrariedade extrema das categorias e o gerativismo, na universalidade extrema, sem, contudo, dar às classes uma relevância maior.[6]

---

6 Podem-se enquadrar as teorias formalistas, como o estruturalismo e o gerativismo, no arcabouço teórico do que Du Bois (1985) chama de "estruturalismo autônomo"; com efeito, ambas as perspectivas restringem severamente a conexão da língua com os fatos positivos do mundo externo, ao considerarem que o

Entendo que a resposta mais pertinente a essa questão é assumir um ponto de vista teórico que, por definição, admita princípios cognitivos, como tem sido extensamente adotado nos trabalhos funcionalistas mais recentes que tratam de categorias gramaticais, incluindo a GDF.

Como a concepção de linguagem que adoto é a de um sistema não autônomo, é possível estabelecer uma relação de intimidade e de dialogicidade entre a própria linguagem, por um lado, e faculdades cognitivas mais gerais, por outro. A possibilidade dessa relação dialógica desloca a linguagem para o contexto da interação do homem com seus semelhantes e com as circunstâncias sociais mais abrangentes.

Nesse enfoque, não há lugar para uma divisão nítida entre faculdades linguísticas e não linguísticas, entre fatos linguísticos e fatos não linguísticos, entre competência e desempenho (cf. Taylor, 1989, p.18).

Além disso, se as categorias devem ser postuladas como detentoras de fronteiras nítidas, e se devem ser definidas em termos da conjunção de traços binários, categóricos, com base em critérios ao mesmo tempo suficientes e necessários, é melhor renunciar antes mesmo de prosseguir. O resultado sempre frustrante é que somente as entidades que exibirem todos os traços da categoria é que são presumivelmente membros dela, sendo excluídas as que não tiverem a mesma sorte.

Ao contrário disso, defendo aqui o princípio de que as categorias contêm membros contínuos sem limites claros entre si. Esse princípio deriva o postulado de que alguns membros de uma categoria compartilham traços comuns que os definem como prototípicos, e outros que não compartilham todos os traços comuns constituem membros marginais.

Isso significa rejeitar a ideia de que a linguagem seja gerada por regras lógicas e por traços semânticos "objetivos", o que implica au-

sistema linguístico é governado por forças inteiramente internas e autônomas. Uma discussão mais detalhada dessa visão será desenvolvida na Seção 2.1.

tomaticamente rejeitar o postulado de que haja um nível estrutural de significação distinto do nível em que o conhecimento do mundo está associado às formas linguísticas. Significa, por fim, rejeitar o princípio da arbitrariedade radical do signo e, por extensão, o postulado de que as categorias são discretas e homogêneas (cf. Silva, 1997).

Para retornar, com outro olhar, à definição nocional, semântica, de categorias gramaticais, a linguística funcional-cognitiva necessita de um tratamento mais sofisticado que o da tradição gramatical, e a alternativa para incluir o que é comum a todos os membros de uma classe gramatical é investir nas definições tradicionais e incorporá-las num tratamento de natureza prototípica.

De acordo com a tradição gramatical, dizer que um nome designa pessoas, lugares e coisas equivale a dizer que um nome designa uma entidade tridimensional concreta e discreta, ou, em termos de Langacker (1987), uma região demarcada no espaço tridimensional. Mediante a projeção do esquema de "coisa" para outros domínios, unidades linguísticas que recortam regiões de domínios não espaciais, como "cor", "tempo", ficam incluídas na categoria dos nomes. Mediante uma relação metafórica entre substâncias concretas e propriedades abstratas, é possível também explicar o caráter *sui generis* de nomes como *amor* e *inveja*. Uma visão prototípica implica afirmar que alguns nomes, certamente os concretos e contáveis, são os exemplos mais representativos da categoria, sendo reservados a outros um caráter marginal.[7]

O enfoque das classes de palavras formando uma escala com os polos ocupados por membros mais prototípicos foi adotado por Givón (1979) e por Hopper e Thompson (1984). Estes autores sustentam a posição de que é possível predizer, em termos universais, que certas percepções de entidades prototipicamente próximas de

7 A importância do conceito de prototipicidade está correlacionada ao comportamento gramatical dos SNs, que podem ser hierarquicamente ordenados de acordo com a acessibilidade a várias regras sintáticas. Os SNs mais acessíveis são os que se referem a entidades animadas, conscientes, volicionalmente ativas, isto é, primariamente seres humanos (cf. Ross, 1973 apud Taylor, 1989, p.192).

uma "coisa" sejam gramaticalmente codificadas como nome, enquanto percepções prototípicas de ações ou eventos sejam gramaticalmente codificadas como verbos.

Com efeito, Givón (1979) postula que o que distingue essencialmente nomes de verbos se sustenta na estabilidade temporal contínua de seus referentes.

Num polo do *continuum*, acham-se entidades com o grau mais elevado de estabilidade temporal e que, por isso, não alteram sua identidade com o tempo, como os membros prototípicos dos nomes.

Membros prototípicos da classe dos verbos, por outro lado, referem-se a entidades sem estabilidade temporal, ou seja, eventos e mudanças rápidas de estado. Na região média do *continuum,* que se vê em (1-1), situam-se as duas outras principais categorias predicadoras, o adjetivo e o advérbio, acompanhados, conforme for o caso, de nomes e de verbos respectivamente.

(1-1)

Mudança rápida	Estados temporários	Estados inerentes permanentes	Maior estabilidade temporal
Verbos	Verbos e Adjetivos	Adjetivos e alguns Nomes	Nome

(Adaptado de Givón, 1979, p.55)

O membro prototípico de uma categoria deve exibir todos os atributos que a representam e nenhum que represente outra categoria. Uma consequência desse postulado é que um membro prototípico da classe dos nomes deve ser maximamente distinto de um membro prototípico da classe dos verbos e o mesmo é válido para outras categorias.

No discurso, em que se introduzem participantes e se relatam eventos, nome e verbo exibem o mais alto grau de contraste morfossintático. Assim, à medida que executar sua função prototípica de introduzir participantes e de desenvolvê-los como suporte do texto, uma forma linguística será codificada como nome e manifestará o

conjunto mais completo possível de constituintes convencionais de que dispõe na língua em questão.

Como um corolário desse enfoque não discreto das classes, para ficar apenas nos dois extremos do *continuum*, pode-se dizer que os nomes devem ser caracterizados com base em sua existência no espaço e, desse modo, o típico referente do nome é uma coisa durável, identificável. Referentes verbais, por outro lado, têm existência somente num certo ponto da continuidade temporal, consistindo num evento identificável.

De certo modo, essa visão de que alguns membros atuam como protótipos da classe toda está presente também no modo como, segundo Câmara Jr. (1972), a língua verbaliza o referente, já que, conforme as palavras do autor, "o nome e o verbo são deduzidos da mesma realidade objetiva ou do mundo real, conforme ela se apresenta ao nosso espírito sob um aspecto ESTÁTICO ou sob um aspecto DINÂMICO" (Câmara Jr., 1972, p.151), para chegar, mais adiante, à conclusão preliminar de que "há uma classificação natural dos semantemas em nominais e verbais; donde vocábulos que são NOMES e vocábulos que são VERBOS, segundo o valor ESTÁTICO ou DINÂMICO do respectivo semantema" (Câmara Jr., 1972, p.153).

Segundo Hopper e Thompson (1984, p.152), essa dualidade perceptual se reflete numa tendência universal que consiste em associar entidades temporalmente estáveis com a classe gramatical dos nomes e entidades não temporalmente estáveis com a classe gramatical dos verbos. Essa correlação não é, contudo, verdadeira na prática, visto que muitos verbos podem denotar situações altamente estáveis, como brilhar aplicável ao sol, por exemplo, enquanto nomes podem denotar situações temporárias, como fogo, ou entidades abstratas, como justiça, que não podem ser diretamente percebidas.

A distinção elaborada por Lyons (1977, p.442-47) entre entidades de diferentes ordens fornece uma solução plausível para essa inconsistência. Nomes comuns, que têm um referente concreto, representam entidades de primeira ordem, enquanto os que se referem a estados de coisas e proposições representam, respectiva-

mente, entidades de segunda e de terceira ordem. Lyons aponta para a existência de uma hierarquia implicativa entre entidades de primeira, de segunda e de terceira ordem. A existência de entidades de terceira ordem implica a existência de entidades de segunda e a existência de unidades de segunda ordem implica necessariamente a existência de entidades de primeira.

Essa escala dá um tom de categorialidade prototípica à subclasse que denota entidades de primeira ordem e está plenamente de acordo com o papel fundamental dos nomes concretos para crianças em fase de aquisição da linguagem, que se vê na tendência para associar nomes com coisas físicas e verbos com ações concretas e visíveis (Brown, 1958, p.247-252).

Apesar de reconhecerem essas correlações, Hopper e Thompson (1984) consideram equivocada essa solução primariamente semântica que Lyons (1977) postula para a distinção entre categorias nominais com base nas três ordens de entidades. Esses autores preterem o critério semântico com base na referência em favor de um critério pragmático com base na função discursiva da linguagem como o fator principal para a motivação da categorialidade. Os traços semânticos parecem constituir condição necessária, mas não suficiente, para atribuir uma dada forma a uma classe lexical.

A prototipicidade de categorias linguísticas não depende somente de propriedades semânticas independentemente verificáveis, mas também e mais crucialmente da função linguística no discurso. Um nome referente a uma entidade prototípica de primeira ordem, como *gato*, não garante a preservação desse traço em todos os usos do radical correspondente.

Isso implica que uma divisão absoluta, não contextual do léxico em radicais nominais e radicais verbais tem uma validade severamente restrita. Um enunciado como (1-2), por exemplo, contém uma instância do radical [*gat*-], que falha em ativar uma entidade concreta, perceptível nesse uso. É em outra instância do mesmo radical, em um enunciado como (1-3), que se atinge o grau extremo de prototipicidade, justamente quando a entidade é referida como um participante no discurso (Hopper; Thompson, 1984, p.155-6).

(1-2) *Gatos bebem leite.*

(1-3) *Ao entrarmos na cozinha, vimos o gato preto da vizinha em cima da mesa bebendo calmamente o leite.*

Os autores defendem a ideia de que as propriedades da semântica lexical são secundárias em relação aos papéis que nomes e verbos exercem no discurso. Nesse caso, os traços semânticos que identificam membros prototípicos das classes dos nomes e dos verbos se derivam dos papéis discursivos que elas exercem. Os autores defendem a hipótese de que o grau de nominidade prototípica correlacionado aos traços morfossintáticos correspondentes mantém uma relação icônica com o grau em que a forma em questão serve para introduzir um participante discreto no discurso.

Quanto ao membro prototípico da classe dos verbos, os traços semânticos poderiam ser similarmente considerados visibilidade, cinese e efetividade, que são justamente os que identificam as percepções atribuídas pelas crianças em processo de aquisição e os que tendem universalmente a realizarem-se mais como verbos.

Similarmente ao que postulam para os nomes, Hopper e Thompson (1984) alegam que os traços semânticos são necessários, mas não suficientes, para identificar o verbo. Assim, para uma forma se qualificar como representante da classe lexical dos verbos, é necessário que represente também a ocorrência de um evento do discurso. Assim, uma forma como *pintando* em (1-4) está muito distante do protótipo, enquanto a forma *pintou* do mesmo item lexical em (1-5) identifica um exemplar prototípico de evento.

(1-4) *O homem pintando a paisagem é um artista famoso.*

(1-5) *Aquele artista famoso pintou o retrato do governador no ano passado.*

Na visão de Hopper e Thompson (1984), o papel discursivo é o principal fator que motiva a identificação de uma forma como mem-

bro central ou periférico de sua categoria. A principal consequência disso é a existência de uma relação icônica direta entre a estrutura morfossintática da categoria e o grau em que ela serve para identificar um participante no discurso, no caso dos nomes, e para relatar um evento discursivo discreto, no caso dos verbos.

Esses autores fornecem, portanto, uma nova perspectiva para enquadrar as categorias do léxico em classes de palavras com base na natureza diagramaticamente icônica delas, ao proporem o Princípio da Iconicidade das Categorias Lexicais, conforme segue abaixo:

> Quanto mais uma forma se referir a uma entidade discursiva discreta do discurso ou descrever um evento discursivo discreto, tanto mais distinta será sua forma linguística em relação a formas vizinhas, tanto de um ponto de vista paradigmático quanto de um ponto de vista sintagmático. (Hopper; Thompson, 1984, p.151)[8]

De uma perspectiva discursiva, nomes introduzem participantes no discurso, que são posteriormente desenvolvidos e manipulados ao longo da constituição do texto. O Princípio da Iconicidade das Categorias Lexicais prediz que, na medida em que uma forma linguística executar essa função, postulada como cardinal, ela será decodificada como nome e manifestará o conjunto possível de determinações nominais que está codificado numa dada língua.

Esse comprometimento com a função discursiva das formas linguísticas assegura que a codificação dos participantes no discurso não depende absolutamente da noção semântica de referencialidade como existência, presente nas semânticas formais de base lógica, como alegam também Du Bois (1980), Givón (1984) e Taylor (1989).

Pelo contrário, a codificação dos participantes depende mais diretamente da intenção comunicativa do falante ao enunciar seu

---

8 Cf. o original: *"The Iconicity of Lexical Categories Principle: The more a form refers to a discrete discourse entity or reports a discrete discourse event, the more distinct will be its linguistic form from neighboring forms, both paradigmatically and syntagmatically".*

discurso; mais especificamente ainda, ela depende do grau de importância para a continuidade discursiva de um argumento particular introduzido previamente, isto é, se é importante a identidade específica ou a identidade genérica da entidade em questão.

É precisamente a especificidade ou a genericidade da identidade do argumento que institui a principal correlação com a alta ou a baixa categorialidade dos nomes, conforme afirma o Princípio de Iconicidade das Categorias Lexicais de Hopper e Thompson (1984).

Du Bois (1980) e Givón (1981) usam o termo *referencial* para entidades relevantes para a continuidade discursiva e o termo *não referencial* para entidades não relevantes. Os dados tipológicos discutidos por Hopper e Thompson (1984) sustentam os postulados de Du Bois (1980) e Givón (1981), mas eles preferem, ainda assim, empregar os termos *manipulável* e *não manipulável* para se referirem respectivamente às categorias que são discursivamente relevantes e às que não são, uma vez que a noção de referencialidade mantém fortes conotações lógico-semânticas herdadas da tradição linguística.

Quanto ao verbo, o Princípio da Iconicidade de Categorias Lexicais prediz que a prototipicidade dessa classe está diretamente correlacionada ao grau em que ela é empregada para relatar eventos. Essa restrição exclui já de início as formas que, em sua *Aktionsart*, exibem propriedades estativas como *saber*, *amar* etc.[9]

Os autores fornecem um conjunto alentado de evidências tipológicas que sustentam o princípio de iconicidade. Essas evidências demonstram que um verbo codifica todos os traços disponíveis na gramática se for plenamente capaz de descrever um evento discursivo real.

Afirmar que formas linguísticas têm uma existência *a priori* fora do discurso, o chamado significado categorial, em contraste com o

9 Um bom argumento para sustentar essa exclusão dos estativos do polo da prototipicidade encontra evidência no fato de que, em muitas línguas, itens lexicais com valor semântico estativo de nomeação de propriedades não são verbos, mas adjetivos (Hopper; Thompson, 1984, p.168).

significado relacional, implica que elas são acategoriais em virtude da irrelevância desse tipo de classificação. A atribuição de categorialidade só tem relevância se a realização, como nome ou como verbo, for imposta à forma pelas motivações discursivas.

Longe de ser aprioristicamente dadas como categorias prontas para a construção de sentenças, nome e verbo manifestam efetivamente seus traços categoriais ao serem requeridos pelo discurso. O que essa propriedade discursiva tem que ver com o princípio de iconicidade entre forma e função?

Quanto menos for requerido pelo discurso (seja para relatar um evento discursivo discreto, seja para introduzir uma entidade discreta para a manipulação potencial na sequência do fluxo discursivo), tanto menos saliente o grau de marcação morfossintática do elemento linguístico como membro da categoria que a linguagem designa universalmente para veicular essas funções. Esse princípio será retomado mais adiante no Capítulo 6.

1.4 As classes predicadoras de palavras

O enfoque escalar das categorias lexicais parece ser o que mais preserva a noção de relatividade preconizada já pelos estruturalistas e o que mais minimiza a aparente universalidade das classes, condição necessária para a adequação tipológica de qualquer enfoque funcionalista. Considerando as classes lexicais, Hengeveld (1992) define as categorias de predicados conforme segue em (1-6).

(1-6) Um predicado *Verbal* é um predicado que, sem a consideração de outros critérios, tem *somente* um uso predicativo na função de núcleo de SV.

Um predicado *Nominal* é um predicado que, sem a consideração de outros critérios, pode ser usado como o núcleo de um SN.

Um predicado *Adjetival* é um predicado que, sem a consideração de outros critérios, pode ser usado como um modificador de um núcleo nominal.

> Um predicado *Adverbial* é um predicado que, sem a consideração de outros critérios, pode ser usado como um modificador de um núcleo não nominal.
>
> (cf. Hengeveld, 1992, p.58)

Como se vê, essas definições são funcionais e, portanto, não nocionais. Elas excluem a possibilidade de um predicado verbal ser usado com uma função não predicadora, mas deixam aberta a possibilidade de predicados nominais, adjetivais e adverbiais serem usados predicativamente. Hengeveld usa a posição das quatro classes acima na estrutura subjacente da oração para fornecer definições funcionais para as categorias lexicais.

Essa posição teórica implica que o escopo dessas definições abriga apenas predicados lexicais básicos, como *homem*, e lexicalmente derivados, como *pintor*, mas, em função da condição restritiva "sem a consideração de outros critérios"[10] (Hengeveld, 1992, p.58), exclui predicados que não dispõem de uma representação lexical na categoria de classe predicadora, ou seja, constituintes sintaticamente derivados, como orações adjetivas, sintagmas preposicionais em função de adjunto adverbial etc.

Para dar um exemplo ilustrativo, em português, o item lexical *frívolo* pode ser usado como modificador mediante uma simples inserção na posição sintagmática adequada. Se, por outro lado, a mesma palavra for empregada na posição de predicativo, haveria a necessidade de uma estratégia morfossintática adaptativa, que exigiria, no caso do português, a inserção de cópula, conforme se observa em (1-7a-b):

(1-7) a *A moça frívola.*
b *A moça é frívola.*

O grau de funcionalidade dessas definições pode ser verificado no próprio comportamento sintático das classes, o que leva Henge-

10 Cf. o original: *"without further measures being taken"*.

veld (1992) a definir *frívola* como um adjetivo. Um enfoque funcional de categorias lexicais com base numa distinção nítida entre unidades lexicais e unidades sintáticas tem vantagens sobre outras abordagens semântica ou morfossintaticamente orientadas (Hengeveld, 1992, p.49-52).

As quatro classes de palavras predicadoras podem ser definidas com referência às categorias funcionais em que elas atuam, conforme definidas em (1-6); assim, verbos (V) são itens lexicais que podem atuar por si mesmos como núcleos de um SV ou uma predicação. Nomes (S) são itens lexicais que podem funcionar por si mesmos como núcleos de um SN. Adjetivos (A) são modificadores de núcleos nominais, enquanto Advérbios de Modo são modificadores de núcleos verbais.

Com base nessas definições funcionais, os advérbios ficam restritos aos de modo, também denominados por Ilari et al. (2002) predicativos qualitativos. O escopo imediato deles é o predicado verbal, principal razão para essa classe de predicados ter justamente recebido o rótulo que tem, ou seja, *advérbio,* embora, como predicadores, também possam tomar o adjetivo em seu escopo. Pode-se dizer que a noção prototípica de advérbio está na afirmação tradicional de que ele modifica tipicamente o sentido do verbo e do adjetivo.

Segundo Ilari et al. (2002, p.73), nessa afirmação está implícita a hipótese de que ele expressa uma espécie de predicação de grau superior: assim como o verbo ou o adjetivo atribuem uma ação ou uma propriedade ao sujeito, o advérbio predicaria uma propriedade da qualidade ou ação que se atribui ao sujeito. Assim, em (1-8) abaixo, o advérbio *calmamente* predicaria uma propriedade da ação de fazer que se atribui ao SN *a gente* como sendo calma.

(1-8) *se a gente for parar para fazer as coisas calmamente não dá* (D2-SP-360)

Já o adjetivo na conceituação acima também fica restrito à classe dos atributivos predicadores, assim rotulados por Müller et al. (2002), que os opõem aos argumentais; as autoras adotam, em

grande medida, algumas propriedades sintáticas atribuídas por Casteleiro (1981) aos predicadores, como admitir paráfrase por oração relativa, admitir tanto *ser* quanto *estar*, poder ser usado como aposto, admitir variação de grau, aceitar tanto anteposição ao núcleo nominal quanto coordenação com outros adjetivos do mesmo tipo (Müller et al., 2002, p.323). Embora alguns desses adjetivos predicadores aceitem anteposição, considera-se aqui prototípico o adjetivo atributivo predicador posposto ao núcleo nominal como o adjetivo *lascada* em (1-9).

(1-9) *o paleolítico é o período... da pedra lascada... como vocês todos sabem* (SP-EF-405)

Hengeveld (1992) evita fazer uso da noção de prototipicidade para enfatizar as funções distintivas das classes de predicados, que ele não vê como necessariamente prototípicas. Para dar uma justificativa, ele menciona um argumento de Thompson (1988 apud Hengeveld, 1992, p.49), que aponta para uma frequência maior de uso de adjetivos na função predicativa, não na função atributiva, que Hengeveld considera a função distintiva dos predicados adjetivos.

Considerar a função modificadora de um núcleo nominal como a função prototípica dos adjetivos pode causar certo constrangimento, já que um postulado desse tipo exclui o uso de adjetivos na função predicativa mediada por cópula. Embora seja uma função predicadora, não é, todavia, exclusiva dos adjetivos, já que nomes podem também ocupar a posição de predicativos numa função também não prototípica da classe dos nomes.

A literatura aponta para a diferença entre adjetivos na função de modificador e adjetivos na função de predicativo. Um mesmo item lexical, como o adjetivo *ciumento*, apresenta propriedades temporárias, menos duráveis na função de predicativo, em (1-10a), e propriedades menos temporárias e mais duráveis na função de modificador, em (1-10b), como se fosse uma propriedade inerente da entidade representada por *namorada* (cf. Bolinger, 1967 apud Hopper; Thompson, 1984).

(1-10) a *Sua namorada é ciumenta.*
b *A namorada ciumenta de meu amigo veio à festa vestida como uma perua.*

Isso explica por que, por exemplo, línguas como o mandarim não dispõem de distinção formal entre verbos e adjetivos em função predicativa. Para o adjetivo poder ser usado atributivamente, necessita, em situações de contraste, de uma espécie de sufixo nominalizador (Hopper; Thompson, 1984).

De um ponto de vista transistêmico, as propriedades que são mais provavelmente lexicalizadas como adjetivos são as mais duráveis, como tamanho, largura, comprimento, cor etc. Já as propriedades menos duráveis e, portanto, mais próximas do conceito de evento são mais provavelmente lexicalizadas como verbos (Dixon, 1977 apud Hopper; Thompson, 1984, p.727).

Embora o adjetivo não possa ser considerado um verbo, a posição de predicativo que ele toma é um traço que o aproxima daquela classe, mas a gramática do português requer o uso de cópula para a marcação número-pessoal e modo-temporal, que não é licenciada para a classe dos adjetivos. Na função modificadora, por outro lado, o adjetivo compartilha com o nome a extensão de propriedades referenciais.

Um critério usado para medir o grau de prototipicidade de um membro de uma classe é o de frequência. Taylor (1989) afirma que o critério de uso tem sido realmente aplicado para explicar por que certos exemplares de uma categoria adquirem o estatuto privilegiado de membros prototípicos num conjunto contínuo e gradiente. Entretanto, rejeita logo esse argumento com base em evidências empíricas cuidadosamente controladas por Rosch (1976 apud Taylor, 1989), que provam que o grau de prototipicidade independe da frequência de ocorrência.

A impressão de uso mais frequente pode ser mais um sintoma do que causa de prototipicidade (Taylor, 1989, p.52). Isso demonstra que mesmo a distribuição sintática, como predicados, das quatro classes de palavras, elaborada por Hengeveld, pode receber uma definição prototípica.

Tanto a linguística funcional da sentença em suas versões conhecidas, como a gramática sistêmica (Halliday, 1985), a Teoria da Gramática Funcional (Dik, 1989, 1997) e as várias tendências funcionalistas norte-americanas (Du Bois, 1985; Hopper; Thompson, 1984; Givón, 1979), quanto a linha funcional-cognitivista (Taylor, 1989; Langacker, 1987) rejeitam a ideia da autonomia da sintaxe em relação à semântica e à pragmática, o que significa argumentar fortemente a favor de uma base semântica e, sobretudo, pragmática para categorias, como nome, verbo, adjetivo e advérbio.

Esse tipo de abordagem será desenvolvido mais adiante, no Capítulo 6 da Parte II, com base na distinção entre nome e verbo e a natureza categorialmente complexa da subcategoria das nominalizações. A próxima etapa, porém, é manter o foco na teoria para caracterizar a perspectiva teórica adotada, com uma argumentação que mostre a evolução do modelo de GDF.

2
Ajustando o foco: uma concepção discursivo-funcional de gramática

Introdução

O tema central e condutor deste trabalho – continuidade funcional das categorias linguísticas e, por conseguinte, o caráter prototípico dos membros mais centrais de cada classe – só é possível no interior de um aparato teórico que questione o princípio da homogeneidade do sistema linguístico e o postulado do caráter discreto das unidades linguísticas. Isso é possível mediante um enfoque que enxergue a linguagem enquanto discurso, e postule a heterogeneidade – incluindo a categorial – como princípio constitutivo do sistema linguístico. Pode-se afirmar que essa perspectiva teórica, que se comprova na formulação de conceitos como variação, sujeito, uso, interação, contexto social, condições históricas e ideológicas, confirma-se como uma das correntes mais fortes dos estudos linguísticos contemporâneos.

O papel deste capítulo é mostrar como um modelo teórico que pretende explicar o funcionamento da gramática de uma língua se encaixa na relação entre sistema e uso, explorando a linguagem a partir do caráter heteróclito de seus fenômenos, um modo de ver que substitui a natureza discreta, descontínua, pela natureza contínua e multifuncional das categorias de análise. Pretendo, por

conseguinte, demonstrar que essa multifuncionalidade se ajusta perfeitamente a modelos teóricos que veem a gramática como um componente seriamente comprometido com a interação entre os níveis pragmático, semântico e morfossintático, como a Teoria da Gramática Funcional de Dik (1989; 1997) e a Gramática Discursivo-Funcional (Hengeveld 2004a; 2004b).

É cara aos modelos funcionalistas a necessidade teórica de tratar aspectos formais da linguagem, vistos não como constituintes de um sistema autônomo, que basta a si próprio, mas sim diretamente vinculado aos fatores mais característicos do discurso, da língua em uso no contexto social, justamente o aspecto da linguagem considerado, pelas teorias formais, como caótico, assistemático e desordenado.

Para demonstrar que o quadro teórico da Gramática Discursivo-Funcional abriga satisfatoriamente o princípio da multifuncionalidade detectável no discurso, prevê-se o desenvolvimento do seguinte roteiro para este capítulo: na Seção 2.1, inicio com uma breve reflexão sobre o recorte epistemológico que isolou o sistema linguístico de seu uso; em seguida, na Seção 2.2, argumento que a Teoria da Gramática Funcional (GF) representa uma das tentativas de superação desse idealismo fundador, mediante a construção de uma gramática funcional e formalizada do discurso, tratando, na Seção 2.3, especificamente do modelo de representação em camadas postulado por Hengeveld (1987) e incorporado por Dik (1989); na Seção 2.4, passo a enfatizar movimentos recentes na GF que reafirmam a necessidade de expansão do modelo em direção ao discurso e, finalmente, faço, na Seção 2.5, uma apresentação do arcabouço geral da Gramática Discursivo-Funcional (GDF), mostrando que sua abrangência ultrapassa o nível intraoracional em direção à dimensão discursivo-interativa.

2.1 O objeto da linguística: idealização e categoricidade

Um dos pontos menos controvertidos na história recente é de pronto a rejeição sistemática da atitude prescritiva da gramática

tradicional pela linguística contemporânea, cuja natureza científica apela para princípios descritivos puros, em que imperam a imparcialidade e a objetividade que Martinet (1972, p.3) tão bem enfatiza. Observe a seguinte citação:

> Como o objeto dessa ciência constitui uma atividade humana, é grande a tentação de abandonar o domínio da observação imparcial para recomendar determinado comportamento, de deixar de notar o que realmente se diz para passar a recomendar o que deve dizer-se. A dificuldade de distinguir a linguística científica da gramática normativa lembra a de extrair da moral uma autêntica ciência dos costumes.

Outro ponto, talvez um pouco mais controvertido, é que, com o advento da linguística descritiva, reconhece-se amplamente o princípio de que a linguagem é um conjunto tão complexo de fenômenos que, "se a estudarmos sob vários aspectos ao mesmo tempo, o objeto da Linguística nos aparecerá como um aglomerado confuso de coisas heteróclitas, sem liame entre si" (Saussure, [1916] 1977, p.16). Foi por isso mesmo que Saussure cunhou a dicotomia fundamental *língua* e *fala*, buscando assim estabelecer um princípio mínimo de homogeneidade e organização que permitisse criar um objeto de estudos bem delimitado e absolutamente independente de outras esferas não linguísticas.

Na esteira de Saussure, é o próprio Martinet (1972) quem formula o princípio epistemológico de que uma teoria científica deve necessariamente efetuar uma seleção entre os elementos da realidade que pretende descrever, tendo por base o paradoxo entre a infinita complexidade dos fenômenos da realidade a ser descrita e a natureza necessariamente finita de qualquer descrição científica. Essa natureza paradoxal da relação entre ciência e realidade parece aumentar tanto em dimensão, quando se trata de descrever e explicar os fenômenos da linguagem, que acaba por pôr em dúvida a possibilidade de delimitar um objeto específico de estudos.

É suficientemente reconhecido que a dificuldade no enfrentamento dessa questão já era francamente admitida por Saussure, ao

alegar que "outras ciências trabalham com objetos dados previamente e que se podem considerar, em seguida, de vários pontos de vista; em nosso campo nada de semelhante ocorre [...]. Bem longe de dizer que o objeto precede o ponto de vista [...] é o ponto de vista que cria o objeto" (Saussure, [1916] 1977, p.15).

Essa necessidade de descrever a linguagem sob um ponto de vista que inclua alguns aspectos do fenômeno e exclua outros aspectos é contraditória com o fragmento acima citado de Martinet (1972). Se o investigador opera uma seleção entre os fenômenos empíricos observáveis, segue-se obviamente que ele não é capaz de apreender "o que realmente se diz", mas o que é autorizado pelo recorte metodológico, isto é, uma parte da realidade.

Ao buscar um objeto de estudos bem delimitado e definido, Saussure idealiza a criatura no próprio ato de criação. A *língua* saussuriana é, em última análise, uma essência que representa exatamente a subordinação do objeto a uma determinada perspectiva metodológica. Ao distinguir a língua da fala, Saussure separa o que é geral e social do que é particular e exclusivamente individual. Esse gesto nítido de idealização, que se completa na noção de sistema de relações, cria um objeto científico apartado da rede de relações sociais que constitui o discurso. Separa, além disso, as propriedades que, por serem essenciais, são internas ao sistema linguístico das que, por serem acessórias e acidentais, são externas ao sistema. Ao fazer essa distinção, Saussure cria um objeto de estudos que considera estritamente linguístico.

A posição de Chomsky, que, em vários aspectos, revolucionou a teoria da linguagem na década de 1960, pouco difere da concepção saussuriana no que se refere exclusivamente à delimitação do objeto. A idealização operada por Saussure se completa com a noção de objeto de estudos desenvolvida por Chomsky (1975, p.83), centrada na competência em oposição ao desempenho, definida na citação seguinte:

> [...] um falante-ouvinte ideal, situado numa comunidade completamente homogênea, que conhece perfeitamente a sua língua

> e que, ao aplicar o seu conhecimento no uso efetivo, não é afetado por condições gramaticalmente irrelevantes, tais como limitações de memória, distrações, desvios de atenção e de interesse, e erros (casuais e característicos).

A atividade real e concreta de linguagem, e que de fato a constitui como fenômeno observável, foi mantida no exterior dos limites do objeto de estudos, sob o nome não mais de fala, mas de desempenho, e seu enfoque só seria possível a partir do esclarecimento das regras da competência. Nesse caso específico, o falante é identificado como o sujeito cartesiano, lógico e universal, e as relações sociais que estabelece com seu interlocutor no processo de interação verbal são desconsideradas.

O estruturalismo, um dos paradigmas mais bem-sucedidos no século XX, desenvolvido a partir das ideias de Saussure, adotou uma versão muito forte do axioma da categoricidade, segundo o qual os dados de base para a análise linguística deveriam ser regularizados de modo a eliminar a variabilidade de fato existente na linguagem (Chambers, 1996). Uma das premissas que dão suporte a esse axioma é que seria necessário abstrair a linguagem da variação do mundo real para torná-la coerente e controlável e o melhor caminho para isso é, segundo Joos (1950, apud Chambers, 1996), aproximar a linguística da matemática, para a qual a inconsistência é, por definição, impossível. Assim, para ele, a linguística se identificaria com um sentido extremo de uma mecânica quântica, que exclui toda a possibilidade de gradação infinitesimal em proveito de unidades unicamente discretas e descontínuas.

A posição de Chomsky, que, desde o início, revolucionou em vários aspectos a teoria da linguagem, mantém a concepção saussuriana no que se refere à redução dos fenômenos a um modelo formal, então denominado *competência*. Dada a variabilidade inerente da linguagem humana, a única possibilidade viável de estudar a competência seria isolá-la do desempenho para garantir distância segura entre o objeto de estudo e seu contexto social.

Ao considerar a linguagem como um conjunto de descrições estruturais de sentenças, o que Chomsky (1980) faz é reduzir o conceito de língua ao de gramática como objeto autônomo. Essa redução se produz principalmente porque o conceito de língua como atividade social é muito vago e em geral tão influenciado por considerações de ordem sociopolítica que acaba por não ter qualquer aplicabilidade para uma teoria rigorosa de linguagem (cf. Chomsky, 1980).

Segundo Givón (1979), o espectro dos fenômenos não categóricos da linguagem tem assombrado a linguística gerativa desde seus primórdios, conforme as marcas que a definem em essência, como os conceitos de gramática, gramaticalidade e competência. Givón (1979) sugere que é somente mediante operações extremas de abstração, simplificação e "higienização" extrema dos dados de base que esses construtos podem elevar-se a primitivos teóricos. O processo de "higienização", a que alude Givón (1979), envolve uma omissão deliberada dos fenômenos não categóricos ou o rebaixamento deles ao estatuto de variação no desempenho. No entanto, entende o autor que a simples consideração de uma pequena amostra de dados extraídos do discurso real torna óbvia a ideia de que os fenômenos não categóricos, que são descartados, têm relevância indiscutível para uma explicação adequada dos fenômenos categóricos da linguagem.

A posição de Givón (1979) representa um esforço concentrado pela busca de soluções teóricas alternativas, o que se deu a partir dos anos 1960, quando novas tendências da linguística contemporânea passaram a questionar a submissão da teoria ao princípio da categoricidade, como uma necessidade epistemológica de criar um objeto homogêneo de estudos. Esse questionamento desemboca na tentativa de resgatar justamente o que, na linguagem, era um obstáculo para Saussure e um critério de pouco rigor metodológico para Chomsky, ou seja, a heterogeneidade como princípio constitutivo da linguagem.

Esse novo valor que se alevanta procura aproximar o homem da linguagem, investindo num sujeito ao mesmo tempo social e étnico,

pragmático e ideológico, retórico e argumentativo, isto é, um ser humano investido de múltiplos papéis sociointeracionais e ideológicos que se refletem necessariamente na organização dos sistemas linguísticos.

No âmbito em que atuam, as tendências não formalistas como o Funcionalismo, a Sociolinguística, a Análise do Discurso, a Análise da Conversação estão unidas no mesmo desejo de superar a parcialidade imposta ao objeto de estudo em razão de sua idealização. Um modo de fazê-lo é substituir a noção de sistema fechado de relações internas pelo tratamento do uso efetivo da língua em situações reais de comunicação.

Como se sabe, Saussure (1977) incluía a Linguística numa ciência mais geral dos signos, que denominou Semiologia, mas seus seguidores reconheceram logo a necessidade de dar à linguística um lugar autônomo em relação às outras disciplinas científicas, emergentes como ela no início do século XX. Apesar dessa tendência, a consequência mais evidente do recorte epistemológico que isolou o sistema do uso foi a abertura de espaços para a constante renovação do objeto de estudos em direção ao uso, como uma espécie de presença obstinada do excluído (Maingueneau, 1990); desse modo, a linguística moderna se defronta, ainda hoje, com os excessos da linguagem que teimosamente transbordam dos limites impostos pelo axioma da categoricidade.

O ponto de partida para os funcionalistas é a visão de que a linguagem é, em primeiro lugar, um instrumento de interação social entre seres humanos, usado com o objetivo principal de estabelecer relações comunicativas entre os usuários (Dik, 1989). Esse traço é central na busca de explicações do modo como são as línguas. Dik (1989) alega que o principal objetivo das línguas naturais é o estabelecimento de comunicação inter-humana, sendo os outros objetivos secundários ou derivados.

Esse postulado é, sabe-se, prontamente negado por Chomsky (1980, p.239), para quem a linguagem humana é um sistema para a expressão do pensamento, independente de estímulos externos e da satisfação de necessidades ou propósitos instrumentais. Para

sustentar sua posição, contrária ao postulado funcionalista de que a comunicação é o objetivo central da linguagem, Chomsky menciona casos de uso não comunicativo, como o exercício de função fática com objetivo de apenas estabelecer contato, ou situações de fala dirigida a uma audiência passiva.

As críticas de Chomsky estão voltadas, como se vê, para uma interpretação muito estreita do conceito de comunicação, que fica limitado, assim, apenas à transmissão de conteúdo proposicional. No entanto, para um funcionalista, a noção de comunicação compreende o conjunto todo de eventos de fala, incluindo o estabelecimento e a manutenção de relações sociais, o que significa, na interpretação de Foley e Van Valin (1984), que a conversação informal engajada no estabelecimento de comunhão fática também deve ser considerada como comunicativa.

Givón (1979) entende que, se a linguagem é um instrumento de comunicação, é um tanto estranho tentar entender sua estrutura sem qualquer referência ao contexto comunicativo e à função comunicativa. Com efeito, sustenta Givón que todo o aparato formal não está presente na linguagem por estar pré-instalado no código genético do organismo, mas por codificar funções comunicativas específicas. Como a finalidade comunicativa das diferentes regras gramaticais não é a mesma, também não são as mesmas as consequências da decisão de segui-las ou quebrá-las. A importância comunicativa de uma regra é maior quando dispensar seu uso resulta numa perda maior de eficiência comunicativa.

Desse modo, o falante-ouvinte não é um usuário de uma gramática autônoma, conforme sugere Chomsky, mas é um indivíduo que faz escolhas comunicativas. Assim, se dispensa uma regra de alta relevância comunicativa, ele o faz para obter algum efeito de sentido de natureza poética, metafórica etc. Por fim, a gramática pode consistir em regras ou estratégias alternativas para o exercício da mesma função ou em funções relativamente similares e, assim, a escolha de uma delas teria consequências comunicativas de pequena importância (Givón, 1979). A interação de regras categóricas com regras variáveis não é absolutamente uma questão de tudo ou

nada, mas de gradiência: estratégias comunicativas praticamente imprescindíveis para a eficiência comunicativa tendem a aparecer como regras categóricas, enquanto estratégias cuja violação altera em grau menor a eficiência comunicativa podem aparecer como regras variáveis com frequência tendendo para menos que 100%.

Como se vê, a posição de Givón, favorável ao enquadramento da linguagem como instrumento de interação social, faz parte de uma disputa transcendental em que a linguística contemporânea está engajada por restringir o enfoque da linguagem à competência gramatical ou por expandi-lo em direção da competência comunicativa. De um ponto de vista teórico, essa disputa equivale a encarar a linguagem respectivamente como um objeto formal ou como um objeto funcional.

Ao falar na diferença entre formalismo e funcionalismo, é necessário rejeitar de pronto a ideia de que as línguas não constituem sistemas estruturados com certa continuidade convencional operando na comunidade de fala, que poderia estar presente numa versão ingênua de funcionalismo, que Du Bois (1985) denomina funcionalismo transparente.

As teorias formalistas se enquadram, por seu lado, no arcabouço teórico do que Du Bois (1985) chama de estruturalismo autônomo, baseado no paradigma estruturalista clássico, ou saussuriano. Essa perspectiva restringe severamente a conexão do signo com os fatos positivos do mundo externo, ao considerá-lo governado por forças inteiramente internas e autônomas do sistema linguístico, onde reinam soberanas as oposições entre formas, relações e funções no interior do sistema linguístico.

O estruturalismo autônomo estabelece uma forte distinção entre a linguística interna ou linguística propriamente dita, que é o estudo da língua, e a linguística externa, que abarca um conjunto diversificado de aspectos mais ligados à esfera da fala.

Entende Du Bois (1985) que essa distinção é útil e necessária por apoiar-se nela para mostrar que certos fenômenos podem ser simultaneamente imotivados a partir do ponto de vista sincrônico e motivados a partir de um ponto de vista metagramatical.

Esse segundo ponto de vista trata as gramáticas como sistemas adaptativos, na medida em que são tanto parcialmente autônomos, e daí *sistemas*, e parcialmente suscetíveis a pressões externas ao sistema, e por isso *adaptativos*. Segundo o autor, essa distinção é frutífera apenas se reconhecermos a existência de motivações em competição e o subsequente desenvolvimento de arcabouço teórico para, inicialmente, descrever e analisar a interação das motivações com contextos específicos e, posteriormente, para prever a resolução da competição entre elas.

Em contraste com essa perspectiva, o funcionalismo transparente assume o princípio de que fatores sintáticos aparentemente autônomos constituem, na realidade, os resultados transparentes dos objetivos funcionais do falante, de modo que não é necessário postular nenhum mecanismo sintático arbitrário para a língua. Em outros termos: as únicas forças que governam a sintaxe são as forças positivas da linguística externa, exemplificadas por Du Bois (1985) como mecanismos de processamento de base biológica, intenções do falante etc.

Na medida em que sejam reconhecidos como teorias, o estruturalismo autônomo e o funcionalismo transparente devem ser entendidos como inerentemente opostos e aparentemente irreconciliáveis; entretanto, estudos recentes em gramaticalização (Heine; Claudi; Hünnemeyer, 1991; Traugott; Heine, 1991; Hopper; Traugott, 1993), que têm posto ênfase nas relações entre discurso e gramática, têm tentado integrar parcialmente esses dois enfoques.

A gramaticalização, pelo menos como foi inicialmente entendida, é um processo de mudança de construções relativamente livres no discurso, cuja forma idiossincrática é motivada somente pelos objetivos do falante em função do evento de fala imediato, para construções relativamente fixas na gramática (cf. Hopper; Traugott, 1993). Na visão de Du Bois (1985, p.347), um tratamento adequado do processo de gramaticalização numa teoria que preveja a relação entre discurso e gramática terá que forçosamente reconhecer que, embora a gramaticalização esteja envolvida com influências dos fenômenos do discurso externos ao sistema, a própria

noção de gramaticalização implica que essas influências interagem com as da língua em si como um sistema gramatical.

Ainda segundo Du Bois (1985), o grau de preparação de que dispomos para aferir o significado da teoria linguística se dá a partir de uma perspectiva que vê as gramáticas como sistemas adaptativos e que, portanto, reconhece, em primeiro lugar, a interação entre forças internas e externas. Quando as forças que se originam de fenômenos externos penetram no domínio da língua, elas se encontram e interagem com as forças internas. Tais fenômenos, que são reconhecidos como intrinsecamente linguísticos, são vistos, na perspectiva de Du Bois, como forças dinâmicas em vez de estruturas fixas, categorias ou entidades de qualquer tipo.

Quando se diz que a linguagem é adaptativa, na medida em que responde a pressões do contexto externo, deve-se de pronto rejeitar a teoria do estruturalismo autônomo, que é limitada na tentativa de explicar a estruturação da linguagem, ao reconhecer apenas forças motivadoras internas ao sistema. Quando se diz, por outro lado, que a linguagem é um sistema, e que, como tal, tem certo grau de continuidade de existência, deve-se rejeitar o arcabouço teórico do funcionalismo transparente, que falha em explicar o fato de que categorias gramaticalizadas são preservadas, por assim dizer, numa forma reificada mais ou menos congelada. Mais importante ainda que isso talvez, ele falha em construir apropriadamente o aspecto interno mais fundamental da gramática, o processo de gramaticalização em si mesmo.

Como alternativa para as versões rigidamente ortodoxas do estruturalismo autônomo e do funcionalismo transparente, Du Bois (1985) propõe a busca de uma síntese entre os dois arcabouços, um novo marco teórico que seja capaz de lidar com a interação de forças externas e internas, assim como com a resolução sistemática da competição em que elas se envolvem. A tarefa metodológica inicial é construir um arcabouço para uma linguística real da fala, que seja capaz de incorporar as descobertas dos estudos modernos do discurso.

A GF, da escola funcionalista holandesa, que tem em Simon Dik (1989; 1997) seu principal mentor, faz parte do esforço por expansão, o que significa estabelecer uma interface sempre desejável entre a configuração formal, sintática, da linguagem e os subsistemas semântico e pragmático; para tanto, enfoca não só as regras que governam a constituição das expressões linguísticas, isto é, a linguagem de um ponto de vista representacional, mas também, e sobretudo, as que governam os padrões de interação verbal nos quais essas expressões são usadas, ou melhor, a linguagem de um ponto de vista interpessoal.

Desde seu início, a GF tem sido uma teoria pragmaticamente orientada e uma enorme quantidade de pesquisa sobre a gramática de diversos tipos de línguas naturais, desenvolvida dentro de seu arcabouço teórico, tem esse tipo de orientação.

A maior parte da pesquisa em GF foi produzida com base nas publicações de Simon Dik (1989; 1997). Foi com a publicação do segundo volume da Teoria da Gramática Funcional (Dik, 1997) que um texto relevante veio a conter um capítulo específico dedicado ao discurso, focalizando, portanto, unidades linguísticas maiores que a oração. Ocioso dizer que a fala em uso tem sido objeto de estudo sistemático fora do âmbito da GF, em áreas como Linguística do Texto, Análise da Conversação, Sociolinguística Interacional, que produziram um acervo muito significativo de trabalhos sobre o assunto. Alguns dos princípios estabelecidos nesses trabalhos estão incorporados no capítulo final de Dik (1997).

Recentemente, Hengeveld (2004a) projetou uma nova arquitetura para a GF, denominada por ele Gramática Discursivo-Funcional, refinada em Hengeveld e Mackenzie (2008). Essa nova arquitetura acomoda o discurso como uma parte integral e muito significativa do modelo. Apesar disso, é a gramática que interessa ao arquiteto desse novo esboço, uma gramática formalizada que incorpore fenômenos inerentes ao processamento do discurso entre os interlocutores, e não exatamente a elaboração de uma nova proposta de análise do discurso.

2. 2 A construção de uma gramática funcional e formalizada do discurso

Um dos traços mais marcantes da teoria linguística é uma refinada divisão social do trabalho em função da natureza extremamente complexa do objeto de estudo. Tem sido comum nos últimos anos que o estudo da gramática seja dividido entre dois enfoques alternativos, o funcional e formal. Uma questão central no debate entre funcionalistas e formalistas é saber se as gramáticas são autônomas em relação aos fatores externos a ela ou se, ao contrário, devem ser explicadas em termos desses fatores externos. Segundo Butler (2003), essa questão requer um tratamento mais rigoroso, porque em geral é grosseiramente simplificada diante da complexidade dos conceitos de autonomia e explanação funcional.

Entre as razões por que o uso dos termos *formal* e *funcional* não pode ser considerado esclarecedor, principalmente quando aplicado aos dois paradigmas alternativos, está a ambiguidade no uso do termo *formal*. Segundo Newmeyer (1998), ele pode referir-se tanto à forma ou estrutura gramatical em oposição ao significado ou uso, quanto à utilização de um sistema notacional para expressar de modo preciso e exato as observações e as generalizações linguísticas. Nuyts (1992) considera essa segunda acepção para afirmar que qualquer intenção de modelar um objeto requer formalização, traço metodológico que não pode ser, assim, restrito às teorias formalistas, já que também os funcionalistas empregam sistemas matemáticos de notação.

A controvérsia em torno dos termos *formal* e *funcional* acaba por ocultar o fato de que diferenças de formalização não passam de mera questão de grau. Com efeito, alguns modelos funcionalistas, como a Gramática Funcional (*Theory of Functional Grammar*) e a Gramática do Papel e da Referência (*Role and Reference Grammar*), consistem em teorias com um elevado grau de formalismo e explicitude notacional. A diferença real não está na utilização de sistemas notacionais, mas no grau de independência ou de autonomia das relações formais ou gramaticais em relação ao significado e ao uso

que os falantes fazem delas no contexto comunicativo. É justamente neste sentido preciso do termo "formal" que a diferença entre as gramáticas formais e as funcionais deixa de ser uma questão relativa para ser uma questão absoluta.

Se, com efeito, a simples consideração de um dos modos como o conceito de autonomia se aplica à linguística – a autonomia da sintaxe –, permite postular que fenômenos sintáticos são essencialmente independentes das funções convencionais de natureza discursiva, pragmática e semântica que esses fenômenos exercem (Croft, 1995).[1]

Croft estabelece ainda uma distinção entre dois aspectos da autonomia que, ao cruzar-se com o sentido de autonomia sintática, abre caminho para adicionar ainda mais confusão ao debate: os conceitos de arbitrariedade e autossuficiência (*self-containedness*). No caso da autonomia da sintaxe, a arbitrariedade representa o fato de que as regras e os elementos sintáticos não poderiam derivar-se de propriedades semânticas ou discursivas, ainda que fossem capazes de predizer corretamente o comportamento sintático; já o conceito de autossuficiência representa o fato de que o subsistema sintático conteria elementos e regras que interagem intimamente entre si sem qualquer tipo de propriedade semântica ou discursiva.

Croft (1995) observa que a autossuficiência é o mais forte dos dois tipos de postulados, na medida em que ele acarreta o de arbitrariedade, e a recíproca não é verdadeira. Mesmo assim, várias combinações logicamente possíveis deriváveis desses dois parâmetros permitem estabelecer diferentes graus de autonomia: seria possível, por exemplo, afirmar que a sintaxe é tanto arbitrária quanto autossuficiente; arbitrária, mas não autossuficiente; ou nem arbitrária nem autossuficiente (Croft, 1995, p.491).

---

1 Croft (1995) menciona ainda a autonomia da gramática e a autonomia da faculdade de linguagem; a primeira diz respeito ao postulado de que a gramática é essencialmente independente de fatores funcionais externos relativos ao uso da linguagem como um instrumento de comunicação em contextos sociais; a segunda, ao postulado de que a capacidade linguística constitui um módulo que é essencialmente independente de outras capacidades cognitivas.

É exatamente em função dessa ausência de fronteiras discretas entre as possibilidades que postular a não autonomia da sintaxe não implica necessariamente a ausência de formalização sistemática. Por esse motivo, Hengeveld (1999) defende a ideia de que os termos *formalismo* e *funcionalismo*, ainda que geralmente aceitos como designações de diferentes enfoques dentro da linguística, não são inteiramente adequados, já que incorporam dois diferentes tipos de oposição.

A primeira oposição consiste na perspectiva básica adotada pelas teorias linguísticas, mediante a qual a gramática pode ser considerada ou como um sistema estrutural autônomo ou como um instrumento de interação social. Teorias com cada uma dessas duas orientações podem ser chamadas autônomas e funcionais, respectivamente (cf. Hengeveld, 1998).

A segunda oposição é de uma natureza completamente diferente. Algumas teorias linguísticas manifestam o objetivo explícito de construir um sistema de representação formal, enquanto, para outros enfoques, esse objetivo não tem relevância; teorias desse tipo podem ser chamadas de formalizadas e não formalizadas, respectivamente. Como esses parâmetros metateóricos de autonomia e formalização caracterizam teorias linguísticas independentes entre si, eles podem ser perfeitamente combinados, como mostra a Figura 1:

	Formalizada	Não formalizada
Autônoma	A	B
Funcional	C	D

(Hengeveld, 1998)

Figura 1 – Funcionalismo e formalismo: autonomia e formalização

A Figura 1 contém quatro diferentes tipos de teorias linguísticas que podem ser distinguidos com base em duas oposições. O tipo A de teoria, formalizado e autônomo, caracteriza o modelo gerativo. O tipo oposto, D, funcional e não formalizado, pode ser exemplificado por enfoques como os defendidos por Halliday (1985) e

Givón (1990). O tipo B, não formalizado e autônomo, apesar de ser logicamente possível, não parece ter nenhum representante significativo; talvez nele se enquadre o estruturalismo saussuriano. O tipo C, formalizado e funcional, é o enfoque representado pela Teoria de Gramática Funcional (Dik 1989; 1997) e pela Gramática do Papel e da Referência (Foley; Van Valin, 1984).

Contrastando a G(D)F com os tipos vizinhos na Figura 1, as abordagens do tipo A e D, a G(D)F compartilha um traço com cada uma delas, mas também tem um traço de que as demais não dispõem. Com as teorias do tipo D, a G(D)F compartilha a orientação funcional, mas delas difere, na medida em que usa um sistema de representação formal que, por isso, força o linguista a desenvolver sistematicamente a teoria mediante a formulação consistente de hipóteses. Como as teorias do tipo A, a G(D)F usa um sistema formal de representação que a torna universalmente aplicável, mas também difere delas por manter uma orientação funcional.

A G(D)F é um exemplo de teoria do tipo C, isto é, é um modelo comprometido com a orientação funcional com o objetivo explícito de construir um sistema de representação formal. Esse sistema de representação formal, caracterizado pela introdução de variáveis, níveis e módulos de representação, está diretamente envolvido com a necessidade de caracterizar a função comunicativa da linguagem, integrando, portanto, aspectos comunicacionais na própria arquitetura do modelo formal (Hengeveld, 1998). O sistema subjacente à construção das expressões linguísticas é um sistema funcional. Por princípio, deve ser estudado no marco das regras, princípios e estratégias que regem seu uso comunicativo natural (Dik, 1997). Segundo Garcia Velasco (2003, p.62), um dos maiores êxitos do programa dikiano se abriga precisamente no fato de haver sido capaz de integrar considerações funcionais e comunicativas na própria arquitetura do modelo sem renunciar a um alto grau de formalização.

Como os formalistas, os funcionalistas também analisam a estrutura gramatical, mas assumem o compromisso de que ela é em grande medida condicionada por fatores derivados da função primordial da linguagem, a de instrumento de interação social. Nesses

termos, pode haver uma correspondência não arbitrária entre forma e função, princípio que é prontamente recusado pelos formalistas que defendem expressamente a autonomia da gramática.

A verdade é que nem mesmo o funcionalista mais radical defenderia a tese de que todos os aspectos sintáticos derivam de aspectos não sintáticos, ou, em outros termos ainda mais radicais, defenderia a tese de que a sintaxe de modo algum teria existência independente do discurso. Nessa perspectiva, que Du Bois (1985) denomina de funcionalismo transparente, pode-se incluir a posição de Garcia (1979), segundo a qual, como a sintaxe não existe por si mesma, as línguas podem ser exaustivamente descritas apenas com referência a princípios comunicativos que subjazem à estrutura do discurso.

Na realidade, o que os funcionalistas defendem é um conceito relativo de não autonomia da sintaxe em relação aos componentes semântico e pragmático, já que reconhecem a existência de fenômenos gramaticais não diretamente deriváveis de aspectos semânticos, pragmáticos ou discursivos.

Givón (1979) defende essa posição moderada ao propor uma retirada estratégica dessa posição extremada, que ele parecia adotar em seu próprio texto, com base no princípio de que há realmente muitos fatos que sustentam a existência de um nível estrutural chamado sintaxe; no entanto, a explicação das propriedades desse nível estrutural requer a referência explanatória a fatores substantivos e não apenas a fatores formais da linguagem.

Por conseguinte, em vez de eliminar um nível independente, formal e autônomo da organização estrutural da linguagem, é preferível estabelecer que a sintaxe é uma entidade dependente, funcionalmente motivada, cujas propriedades formais refletem as propriedades dos parâmetros explanatórios que motivam seu surgimento.

A verdadeira diferença entre formalistas e funcionalistas em relação ao conceito de autonomia se centra, segundo Garcia Velasco (2003), não tanto na autonomia da sintaxe, mas na autonomia da gramática, entendida pelos formalistas como a faculdade de linguagem, existente na mente do falante/ouvinte. A sintaxe é assim

autossuficiente no sentido de que abriga princípios independentes, próprios, cuja manifestação física mais evidente é a hipótese de uma faculdade inata de linguagem. Segundo Noonan (1998), isso indica que as categorias gramaticais e suas relações no interior do sistema têm existência independente e anterior à manifestação das expressões linguísticas no contexto de sua produção e justamente por isso podem ser descritas sem referência a esses fatores.

Croft (1995) menciona duas leituras para o conceito de autonomia, a autonomia da sintaxe em relação à semântica e à pragmática, e a autonomia do conhecimento gramatical em relação à orientação social e comunicativa do sistema; como se verá desses dois sentidos relacionados, é possível afirmar que o modelo de Gramática Funcional se encaminha para uma nova concepção da gramática, organizada em módulos, que dá interdependência e, ao mesmo tempo, autonomia aos componentes do Nível Interpessoal ou pragmático, do Nível Representacional ou semântico e do Nível Morfossintático e do Nível Fonológico.

2.3 Níveis e camadas de organização: a Teoria da Gramática Funcional

Como é possível entender um modelo de análise que é, simultaneamente, formal e funcional? Do ponto de vista formal, a GF aplica de modo recursivo a notação lógica tradicional de análise de predicados de um lugar do tipo (x_i: f (x_i)), em que se lê "um x_i tal que a x_i se aplica a propriedade f", e de predicados de dois lugares uma expressão como (x_i: f (x_i) (x_j)), em que se lê "um x_i tal que x_i tem uma relação f com x_j" (Dik 1989, p.63). Considere, por exemplo, a sentença contida em (2-1).

(2-1) *O garoto esperto estava lendo um livro interessante.*

A GF aplica a (2-1) os esquemas de termo contidos em (2-2), para formar os termos de (2-3), que são, por sua vez, aplicados à

predicação; esta, finalmente, representa um estado de coisas, conforme (2-4) sem maiores detalhes formais, como a indicação de operadores de definitude, gênero e número para os nomes e modo-temporal e número-pessoal para os verbos.

(2-2) $garoto_N\ (x_i)_{\emptyset}$
$livro_N\ (x_j)_{\emptyset}$
$esperto_A\ (x_i)_{\emptyset}$
$interessante_A\ (x_j)_{\emptyset}$
$ler_V (x_1)_{Ag}\ (x_2)_{Pac}$

(2-3) $(x_i$: $garoto_N\ (x_i)_{\emptyset}$: $esperto_A\ (x_i)_{\emptyset})$
(x_j) $livro_N\ (x_j)_{\emptyset}$: $interessante_A\ (x_j)_{\emptyset})$

(2-4) $ler_V (x_i$: $garoto_N\ (x_i)_{\emptyset}$: $esperto_A\ (x_i)_{\emptyset})_{Ag})$ (x_j) $livro_N\ (x_j)_{\emptyset}$: $interessante_A\ (x_j)_{\emptyset})_{Pac})$

Para fazer justiça às propriedades formais e semânticas de orações de um modo tipologicamente adequado, Dik (1989) assume que cada oração deve ser descrita em termos de uma representação subjacente (doravante RS). Essa representação se projeta sobre a forma real da expressão linguística correspondente mediante um sistema de regras de expressão que, por sua vez, determina a forma, a ordem e o padrão entonacional dos constituintes em função do estatuto deles na RS da oração:

REPRESENTAÇÃO SUBJACENTE DA ORAÇÃO

↓

REGRAS DE EXPRESSÃO

↓

EXPRESSÕES LINGUÍSTICAS

(Adaptado de Dik, 1989, p.46)

A RS da oração é uma estrutura abstrata complexa em que se devem distinguir diversos níveis ou camadas de organização formal e semântica, que podem ser representados em (2-5):

(2-5) ORAÇÃO → "ato de fala"
PROPOSIÇÃO → "fato possível"
PREDICAÇÃO → "estado de coisas"
PREDICADO → "propriedade/relação"
aplicados a TERMOS → "entidade/entidades"

(cf. Dik 1989, p.46)

A construção da RS de uma oração requer, em primeiro lugar, um predicado que se aplique a um número apropriado de termos. Predicados designam propriedades ou relações, enquanto termos são usados para referir-se a entidades. Como um exemplo de predicado, considere-se o verbo *dar*. Esse predicado designa uma relação entre quatro entidades nos papéis de doador, de algo dado, de um recebedor e de um local, que são aplicados a quatro termos respectivamente, como por exemplo, *João, o livro, Maria* e *biblioteca*. Quando um predicado se aplica a um conjunto apropriado de termos, o resultado é uma predicação, que pode ser provisoriamente representada em (2-6).

(2-6) dar (João) (o livro) (a Maria) (na biblioteca)

(Adaptado de Dik, 1989, p.48)

A predicação designa um estado de coisas, ou seja, algo que ocorre no mundo real ou que pode ser criado no mundo mental de E (emissor) e de D (destinatário); localiza-se no tempo e no espaço; pode ter certa duração; pode ser visto, ouvido ou percebido de alguma forma. Assim (2-6) pode ser representado como (2-7).

(2-7) Pass [[dar(João)(o livro)(a Maria)] (na biblioteca)]]

(cf. Dik, 1989, p.47)

Um elemento como *Pass* representa um meio gramatical destinado a localizar o estado de coisas em um intervalo de tempo que precede o momento da fala (t_o). Esse elemento é chamado de operador de predicação. Um constituinte como *na biblioteca* representa, por seu lado, um meio lexical para localizar, no espaço, o estado de coisas designado pela predicação. Do ponto de vista da configuração interna, esse constituinte tem uma estrutura de termo. Os termos que são requeridos pela semântica do predicado, como (João), (livro) e (Maria), são denominados argumentos do predicado; os termos que fornecem outras informações, tais como (na livraria), são chamados satélites.

Considere agora a seguinte forma de expressão linguística em (2-8) e a RS correspondente em (2-9):

(2-8) *Pedro viu que João deu o livro a Maria na biblioteca.*

(2-9) Pass [ver(Pedro)(e_i)]
e_i = Pass [[dar(João)(o livro)(a Maria)] (na biblioteca)]

(cf. Dik, 1989,p.47)

A predicação toda representada em (2-8) agora aparece como especificação de um novo estado de coisas, ou seja, como segundo argumento de *ver,* que tem em (Pedro) o primeiro argumento. Predicações desse tipo, que podem ser encaixadas na posição de argumento de predicados hierarquicamente mais altos, são chamadas predicações encaixadas, e a predicação mais alta, predicação matriz.

Considere-se agora uma expressão linguística como (2-10):

(2-10) *Pedro acredita que João deu o livro a Maria na biblioteca.*

Embora a oração subordinada de (2-10) tenha a mesma forma que a de (2-9), seu estatuto semântico é diferente. As coisas em que

se pode acreditar não constituem estados de coisas, mas proposições, conteúdos proposicionais ou fatos possíveis. Assim a predicação, que designa um estado de coisas, pode ser construída dentro de uma estrutura de ordem mais alta, a proposição, que designa um conteúdo proposicional, um fato possível. Assim, (2-10) terá a representação contida em (2-11), em que X_i designa uma proposição encaixada:

(2-11) Pass[acreditar(Pedro)(X_i)]
X_i = Pass[[dar(João)(o livro)(a Maria)](na biblioteca)]

(cf. Dik, 1989, p.48)

A análise contida em (2-11) mostra que há um estado de coisas que consiste numa relação de doação entre três entidades representadas por (*João*), (*livro*) e (*Maria*), e que o estado de coisas inteiro está localizado no tempo por um operador de predicação *Passado* e no espaço por um satélite Locativo (*na biblioteca*).

Essas informações são relevantes, mas ainda não dão conta da análise cabal de (2-10), que deve conter também o estatuto pragmático da sentença, determinado pelo ato de fala que ela manifesta, ou pela força ilocucionária de (2-10), que é uma sentença declarativa.

A força ilocucionária se aplica a proposições, não a predicações, já que o objeto de uma declaração, de uma pergunta ou de uma ordem é um conteúdo proposicional e não um estado de coisas, conforme se observa em (2-12), em que DECL indica a forma que deve assumir a força ilocucionária da proposição:

(2-12) DECL(X)
X_i = Pass[[dar(João)(o livro)(ao livreiro)](na livraria)]

(cf. Dik, 1989, p.49)

Uma predicação pode ser construída dentro de uma proposição e esta, por sua vez, pode ser construída dentro de um esquema

ilocucionário; o resultado é a RS da oração plena, que designa um ato de fala (nesse caso, um ato de fala declarativo) com respeito à proposição; esta contém uma predicação que é por sua vez construída a partir de um predicado mais um conjunto apropriado de termos.

Até meados de 1980, a estrutura subjacente era monoestratificada, como sugerido em (2-5). Para mostrar recusa por transformações, a RS da oração era dotada de três componentes básicos: constituintes, categorias e funções. Por constituinte, entende-se um item lexical, como nome, por exemplo; por categoria, um sintagma nominal; e por função, relações sintáticas como sujeito, objeto etc. e relações semânticas como agente, paciente etc. Para explicar como seria possível selecionar as funções necessárias para descrever a linguagem, Dik fornece a seguinte estipulação, antecipando, segundo Anstey (2004), em muitos aspectos a filosofia central que percorre a teoria até aqui: "Existem no mínimo tantas funções em uma língua quanto são as diferenças gramaticais (não semânticas) entre cada expressão que não podem ser correlacionadas a diferenças de constituência e/ou categorização" (Dik, 1968, p.176 apud Anstey, 2004, p.28).[2]

Em 1987, inspirado principalmente por Foley e Van Valin (1984), Hengeveld apresenta pela primeira vez uma RS da oração organizada em três camadas. A oração, agora, não é apenas organizada em camadas na perspectiva horizontal, mas ela é também verticalmente organizada, representando, em dois níveis, dois diferentes aspectos da linguagem, o Interpessoal e o Representacional. A Figura 2 é uma ilustração da organização em camadas da estrutura subjacente da oração proposta por Hengeveld (1987), que foi absorvida por Dik (1989).

---

2 Cf. o original: *"At least so many functions are set up for a language as there are grammatical (not semantic) differences between the linguistic expressions of that language which cannot be correlated to differences in constituency and/or in categorization".*

$(\pi_4 E_1: [(\pi_3 X_1: [(\pi_2 e_1 :[(\pi_1 f_1)\ (x_1)](e_1))]\ (X_1))]\ (E_1))$	oração[3]
$(\pi_3 X_1: [(\pi_2 e_1 : [(\pi_1 f_1)\ (x_1)](e_1))]\ (X_1))$	proposição
$(\pi_2 e_i : [(\pi_1 f_1)\ (x_1)](e_1))$	predicação
$(\pi_1 f_1)\ (x_1)$	predicado

Camada	Variável	Designação	Nível
Oração	(E_1)	Ato de fala	
Proposição	(X_1)	Conteúdo proposicional	Interpessoal
Predicação	(e_1)	Estado de coisas	
Predicado	(f_1)	Propriedade/relação	
Termo	(x_1)	Indivíduo	Representacional

(cf. Hengeveld, 1987)

Figura 2 – A representação dos enunciados na GF

Cada camada na Figura 2 representa diferentes entidades, que podem ser definidas em termos de uma versão estendida da classificação para os tipos de entidades, fornecida por Lyons (1977, p.442-7).

Assim, um predicado (f_1) representa uma entidade de zero ordem, que não tem existência independente, podendo ser avaliada em termos de sua aplicabilidade a outros tipos de entidade. Um termo (x_1) representa um indivíduo, uma entidade de primeira ordem, que pode ser localizada no espaço e pode ser avaliada em termos de sua existência. Uma predicação (e_1) representa um estado de coisas, uma entidade de segunda ordem, que pode ser localizada no tempo e no espaço e pode ser avaliada em termos de sua realidade. Uma proposição (X_1) representa um conteúdo proposicional, uma entidade de terceira ordem, que não pode ser localizada nem no tempo nem no espaço, mas pode ser avaliada em termos de sua verdade. Por fim, a oração (E_1) representa um ato de fala, uma entidade de quarta ordem, que se localiza no tempo e no espaço e pode ser avaliada em termos de suas condições de felicidade (Hengeveld, 1987).

3 A letra grega π representa a inserção de operadores, e os números subscritos, a camada em que são inseridos: da mais externa (de nível 4) para a mais interna (de nível 1).

A estrutura básica da predicação nuclear é determinada pelas possibilidades combinatórias do predicado; uma boa ilustração dessas relações está contida em (2-13), que representa a RS de um predicado trivalente como o verbo *dar*:

(2-13) $dar_V (x_1: \text{<anim>} (x_1))_{Agente} (x_2)_{Meta} (x_3: \text{<anim>} (x_3))_{Recipiente}$

(Adaptado de Dik, 1989, p.68)

O esquema de predicado em (2-13) especifica (i) a forma do predicado codificada sob um tipo padrão de representação fonológica e, no caso de não ser necessário esse tipo de detalhamento, convencionou-se uma notação simplificada com o uso da forma escrita do infinitivo, quando se tratar, como é o caso de (2-13), de um predicado verbal. O esquema de predicado inclui a categoria sintática, que é V nesse caso. Além de verbais, a GF reconhece a existência de predicados nominais e adjetivais.[4] Algumas subcategorias, como a de predicado bitransitivo, são inferíveis com base nas propriedades estruturais do predicado em si mesmo como, por exemplo, o predicado *dar* admitir três argumentos.

Outras informações que o esquema de predicado representa dão conta da valência quantitativa e qualitativa. A valência quantitativa, que se refere ao número de argumentos do predicado com base na predicação nuclear, é configurada em (2-13) pelas variáveis x_1, x_2, x_3, que simbolizam as posições argumentais ocupadas por entidades individuais. Distinguem-se, portanto, predicados de zero lugar ou avalentes, predicados de um lugar, ou monovalentes, predicados de dois lugares, ou bivalentes, e predicados de três lugares, ou trivalentes.[5]

---

4 Os desenvolvimentos mais recentes da GDF reconhecem a necessidade de incluir certos usos das preposições ora na classe dos operadores (preposições gramaticais) ora na classe dos itens lexicais (preposições lexicais); nessa última condição elas atuam como predicados (v. a esse propósito Keizer, 2004b).

5 Outros autores como Borba (1996) consideram, contudo, que predicados de transferência devem ser categorizados como de quatro lugares ou tetravalentes. Na visão de Dik (1989), todavia, Origem e Direção não constituem necessariamente funções semânticas de constituintes argumentais nesse tipo

A valência qualitativa representa a grade temática do predicado, especificada em (2-13) pelas funções semânticas de Agente, Meta e Recipiente e por restrições selecionais, como a do traço semântico *animado*, que se impõe sobre as funções de Agente e Recipiente.

Os predicados podem ser básicos ou derivados. Predicados derivados são produzidos mediante regras de formação de predicados, que nada mais são senão regras de produtividade do léxico, não exatamente regras de transformação da gramática, como as previstas pela versão padrão da Gramática Gerativa, que são, por princípio, rejeitadas pela GF. Os predicados básicos e derivados constituem o que Dik denomina Fundo no esboço geral do modelo da GF, que aparece na Figura 3 a seguir.

Os vários níveis que constituem a RS consistem na formatação da estrutura da oração, que recebe, no final, uma forma de expressão determinada por regras específicas.

A outra parte do Fundo, que consiste em estruturas de termo, Dik (1989) reserva os seguintes postulados: (i) analisam-se como termos as expressões linguísticas usadas para se referir a entidades de algum mundo; (ii) por termo entende-se não apenas itens simples, como pronomes e nomes próprios, mas também sintagmas nominais complexos como *O livro que João comprou* em sentenças como *O livro que João comprou é uma obra-prima de Machado de Assis*.

Estruturas de predicado e estruturas de termo podem ser consideradas básicas e derivadas. São básicas as estruturas de termo, que representam itens simples, como pronomes e nomes próprios, que somente podem ser usados como termos; entretanto, alega Dik (1989) que a grande maioria das estruturas de termo é derivada por meio de regras produtivas de formação de acordo com o seguinte esquema geral:

(2-14) $('\Omega x_i\ (x_i: \Phi_1(x_i): \Phi_2\ (x_i): \ldots : \Phi_n\ (x_i))$

(cf. Dik, 1989, p.55)

---

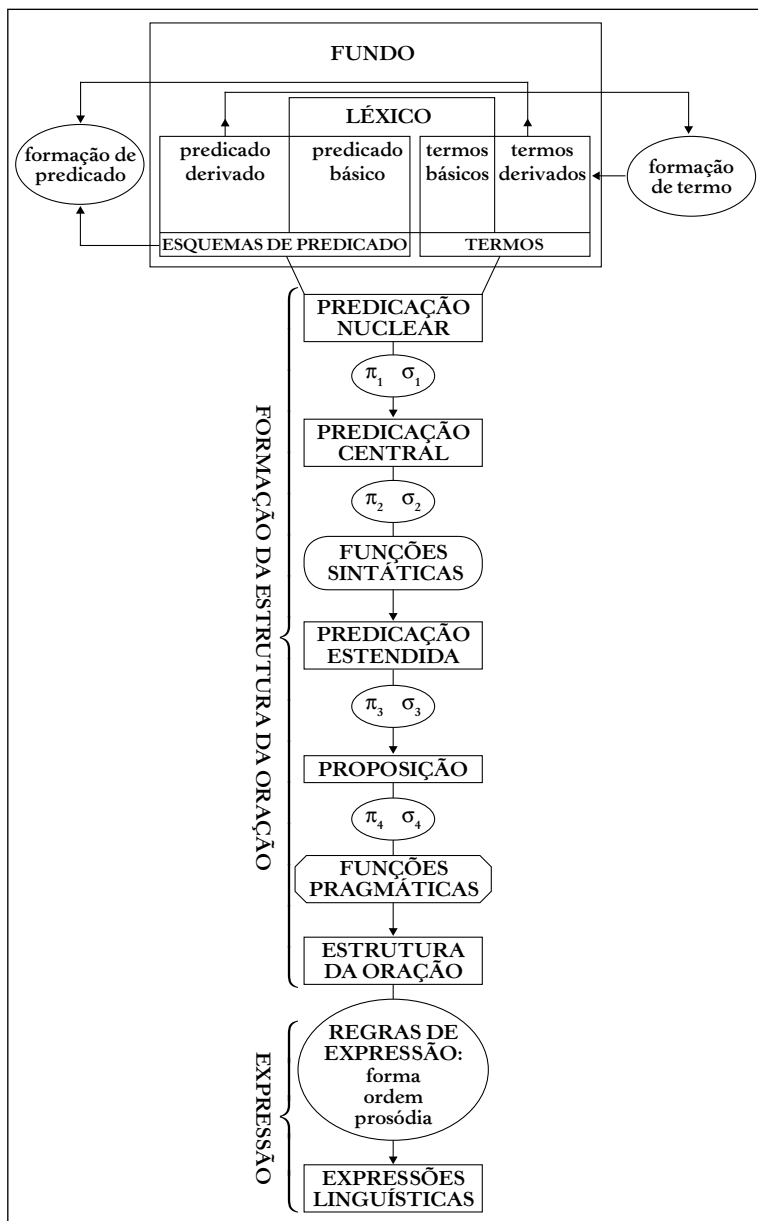
semântico de predicados, mas predicados de movimento, quando requerem um argumento com essas funções semânticas.

(Dik, 1989, p.53)

Figura 3 – Esboço do modelo de representação linguística da GF

No esquema em (2-14), Ω representa um ou mais operadores de termos, x_i simboliza o referente pretendido e cada $\Phi_1(x_i)$ representa "uma predicação aberta em x_i", ou seja, um esquema de predicado em que todas as posições, com exceção daquela ocupada por x_i... x_n, tenha sido preenchida por estruturas de termo. Uma predicação aberta funciona como um "restritor" dos valores possíveis que x_i pode ter. Num SN complexo como *o elefante branco que o juiz construiu,* o primeiro restritor é o esquema de predicado para *elefante*; o segundo restritor é o esquema de predicado para *branco* e o terceiro é construído a partir do esquema de predicado de *construir* mediante a inserção das estruturas de termo para *juiz* na primeira posição argumental do predicado verbal e de um pronome relativo na segunda.

Predicados básicos, armazenados no léxico, representam as propriedades e relações mais fundamentais que são passíveis de receber expressão numa língua. Esses predicados, que, juntamente com suas propriedades combinatórias, são codificados nos esquemas de predicado, têm que ser separadamente aprendidos e memorizados para que o falante seja capaz de usá-los. Em razão de limites cognitivos óbvios quanto ao número de predicados básicos que uma língua pode conter, são as regras de formação de predicado que garantem a projeção dessas propriedades/relações básicas num conjunto muito maior de propriedades/relações (Dik, 1997, p.19).

Uma vez que a existência de predicados está condicionada à de estruturas de predicado, as regras de formação permitem derivar produtivamente novas estruturas de predicado a partir das que já existem. Como projetam estruturas de predicado em outras, as regras de formação tomam uma estrutura como *input* e produzem um esquema de predicado como *output.* O fato de qualquer propriedade do *input* ser passível de modificação mediante uma regra de formação de predicado significa que as várias propriedades codificadas no esquema de predicado fornecem uma base natural para distinguir diferentes tipos de processos de formação.

Assim, segundo Dik (1997, p.6), é possível encontrar processos que afetam cada uma das propriedades de um esquema de predica-

do, como forma, tipo e valência; além disso, uma regra complexa pode afetar diversos traços simultaneamente. Esses processos de formação fornecem uma explicação adequada para modificações na valência quantitativa, traduzíveis como extensão e redução de valência. A extensão de valência envolve, por um lado, os casos em que um predicado intransitivo se transforma num transitivo, como ocorre com a formação comum de predicado causativo. O processo de redução de valência envolve, por outro lado, os casos de incorporação de objeto, redução do primeiro argumento e redução do segundo argumento.

De acordo com Dik, esses processos produzem reflexos evidentes no âmbito da valência qualitativa, já que a alteração de valência quantitativa implica necessariamente alterações das funções semânticas e das restrições de seleção do predicado *input*. Assim, por exemplo, a formação causativa requer o acréscimo de um segundo argumento na função semântica de Causado. Num processo de redução de valência, como o que projeta predicados intransitivos a partir de transitivos, um argumento com função original de Meta (Paciente) pode passar a Processado, no caso de predicado dinâmico e não controlado, ou a Zero, no caso de predicado não dinâmico e não controlado.

Voltando à organização em camadas, vimos que a representação na Figura 1 acima não é mais monoestratificada, embora continue evitando transformações. Ela fornece uma representação separada de um ato de fala (E_1), dentro do qual se comunica um conteúdo proposicional (X_1). O nível mais alto da estrutura é chamado, de acordo com Halliday (1970, p.325), de Nível Interpessoal. Ele é estruturado com mais detalhes em (2-15):

(2-15) (E_1 [(ILL (S) (A) (X_1: [-----] (X_1))] (E_1))

A RS é estruturada em (2-15) com base em um predicado ilocucionário abstrato (ILL), que indica o tipo de ato de fala enunciado e que tem o falante (S), o ouvinte (A) e o conteúdo comunicado (X_1) como seus argumentos.

O conteúdo proposicional, por sua vez, a última camada do Nível Interpessoal, contém a descrição de um estado de coisas (e_1) que envolve uma relação entre um predicado (f_1) entre um ou mais indivíduos ($x_1...x_n$). Essas camadas constituem o nível mais baixo, chamado de Representacional.

Cada uma das unidades relevantes da estrutura da oração pode ainda ser modificada por operadores, que consistem em elementos abstratos representando distinções semânticas e pragmáticas expressas por meios gramaticais e, conforme já mencionado, por satélites, que consistem em distinções semânticas expressas por meios lexicais opcionalmente selecionados. No esquema em (2-16a-b), os operadores são representados por π e os satélites por σ; o número subscrito indica a camada, aqui restrita apenas às do Nível Representacional.

(2-16) a $\pi_3 p_1$: $[(\pi_2 e_1 : [(\pi_1 f_1) (x_1)](e_1))] (p_1))$
b $(p_1$: $[e_1$: $[(f_1$:... $f_1)$: $\sigma_1 (f_1)))(x_1)](e_1)$: $\sigma_2(e_1))] (p_1)$: $\sigma_3 (p_1))$

Os operadores, que fazem parte do sistema gramatical da língua, representam as seguintes distinções semânticas e pragmáticas:

(2-17)	π_4	operadores de ilocução	Estratégia comunicativa do falante (ex.: reforço)
	π_3	operadores de proposição	Validade do conteúdo proposicional (ex.: evidencialidade)
	π_2	operadores de predicação	Cenário do estado de coisas (ex: Tempo, Realidade)
	π_1	operadores de predicado	Mudança na estrutura interna do estado de coisas (ex: aspecto)

(Adaptado de Hengeveld, 1999, p.4)

Todos os tipos de operadores em (2-17) têm funções que são próprias do nível em que operam. Operadores de predicado representam expressões gramaticais que especificam propriedades

adicionais do estado de coisas. Muitas expressões aspectuais são capturadas por esse tipo de operador. Operadores de predicação representam expressões gramaticais que especificam o cenário do estado de coisas, cuidando de representar, por exemplo, a expressão de tempo e de aspecto. Operadores de proposição representam expressões gramaticais que especificam a atitude proposicional do falante, como no caso da modalidade evidencial. Por fim, operadores de ilocução representam expressões gramaticais que modificam a força do ato de fala, cuidando, por exemplo, do uso de morfemas de ênfase.

Essa caracterização semântica encontra um reflexo formal no modo como expressões de Tempo/Aspecto/Modalidade (doravante TAM) são realizadas translinguisticamente, com base na generalização tipológica contida em (2-18):

(2-18) Na medida em que as línguas usam uma única e mesma estratégia morfológica para a expressão de operadores, esses são expressos na seguinte ordem em relação ao predicado: $Pred_V\ \pi_1 > \pi_2 > \pi_3 > \pi_4$ ou $\pi_4 > \pi_3 > \pi_2 > \pi_1$ Pred

(Hengeveld, 1999, p.5)

A organização em camadas pode ser comprovada pelas relações de escopo dos operadores e dos satélites. Considere-se a ordem dos morfemas para a mesma sentença no hidatsa e no inglês em (2-19) e (2-20) respectivamente.

(2-19) *Hidatsa*

$Pred_V$	π_1		π_2	π_3	
Wíra	*i*	*appari*	*ki*	*stao*	*ski*
Árvore	ela	crescer	INGR	PASS	CERT

'A árvore deve ter começado a crescer'[6]

6 PASS (ado) representa a categoria de tempo; INGR(essivo), a de aspecto; CERT(eza), a de modalidade.

(2-20) *Inglês*

		π_3	π_2	π_1	$Pred_V$
The	tree	must	have	begun	to grow.
		CERT	PASS	INGR	crescer

'A árvore deve ter começado a crescer'

(Adaptado de Hengeveld, 1999, p.5)

De acordo com o Princípio de Orientação Centrípeta (Dik, 1989, p.401), a ordenação é determinada pela relativa distância que os constituintes mantêm com o núcleo, o que pode produzir uma organização que simula uma imagem ao espelho em torno do núcleo de acordo com o escopo dos constituintes.

Com base nesse princípio, Hengeveld (1989) já propunha que a ordem dos morfemas em torno do núcleo deve refletir as relações de escopo de modo que 1 representa a posição para aspecto e modalidade orientada para o participante, como a facultativa, a posição 2 representa relações de tempo, realis/irrealis e quantificação, enquanto a posição 3 representa evidencialidade. Os exemplos em (2-19) e (2-20) ilustram essas relações de escopo e, por conseguinte, o princípio de que a sentença deve ser organizada em camadas estruturais. A ordem dos satélites da sentença (2-21a) do inglês, um pouco mais rígida que a do português, consiste também num bom exemplo das relações de escopo (2-21b) e numa comprovação a mais da necessidade da organização formal em camadas postulada pela GF:

(2-21) a *She certainly* (σ_3) *danced beautifully* (σ_1) *recently* (σ_3)
b σ_3 predicado σ_1 σ_2

Esses exemplos representam um fenômeno linguístico universal, que foi primeiramente investigado extensivamente por Bybee (1985). A autora notou que, dentro do domínio de expressões de TAM, o Aspecto tende a ser expresso mais próximo do predicado, seguido primeiramente por Tempo e depois por Modalidade. O tratamento desse tipo de operador pela GF fornece uma base for-

malizada para descobertas como essa mediante a representação da oração por uma organização em camadas.

Importam as funções pragmáticas apenas na medida em que recebam algum tipo de manifestação formal no sistema gramatical da língua em questão, sendo, por isso, passíveis de formalização nos diferentes níveis de representação. Para mencionar um exemplo simples, o português mantém a expressão de um ato ilocucionário exortativo, como em (2-22a), uma sentença mais adequada ao estilo formal e (2-22b) mais adequada ao estilo informal.

(2-22) a *Caminhemos juntos para não nos perder.*
b *Vamos caminhar juntos para não nos perder.*

Essa função, todavia, é formalmente exercida pelo imperativo na 3ª pessoa do plural; nesse caso, teria sentido mencionar o ato de fala exortativo na gramática da língua portuguesa, como, por exemplo, se observa na gramática do tucano; o imperativo é caracterizado pelo acréscimo do sufixo **–ya** em (2-23a), distinguindo esse tipo de ato de fala do exortativo, pelo acréscimo do sufixo **–râ** em (2-23b), e do suplicativo, pelo acréscimo do sufixo **–ma** em (2-23c).

(2-23) a *apê-ya!*
play-IMP.2PL
'Play!'
b *apê-râ*
go-IMP.HORT.1PL
'Let's play!'
c *apê-ma*
play IMP.SUPPL.
'Leave me play!'

Tucano (Ramirez, 1997, p.144-5)

Não há, por outro lado, qualquer marca formal para o suplicativo em português, que, por referir-se apelativamente a uma 2ª pessoa, usa a forma do imperativo e com ele se confunde formalmente.

Não teria sentido, portanto, incluir um o ato de fala suplicativo na gramática do português.

O fato é que, apesar de defender veementemente uma metodologia funcional para o enfoque da linguagem, os adeptos da GF não são, como se vê, refratários a uma formalização dos fenômenos linguísticos. Por causa dessa posição nitidamente intermediária, a GF recebe críticas tanto de formalistas quando de funcionalistas. Os formalistas alegam, por um lado, que a GF não passa de uma variante notacional da Gramática Gerativa; por outro lado, funcionalistas de outras vertentes como Croft (1995) afirmam que a GF está a meio caminho entre o formalismo e o funcionalismo por não reduzir todos os aspectos sintáticos a princípios funcionais com base na preservação de certa autonomia dos níveis pragmático, semântico e sintático.

Há manifestações indubitavelmente programáticas como a que se lê em Dik (1989, p.8), segundo a qual a semântica é instrumental em relação à pragmática, e a sintaxe, em relação à semântica. Sob essa perspectiva, Dik (1989) parece contradizer a análise de Croft (1995), já que não haveria qualquer espaço para uma sintaxe absolutamente autônoma: pelo contrário, na medida em que é possível estabelecer uma divisão clara entre sintaxe e semântica, a sintaxe permite aos indivíduos construir expressões complexas para transmitir significados complexos que permitem aos indivíduos comunicar-se de modo diversificado em situações específicas de interação.

Para dar ainda maior suporte teórico a esses princípios, Dik (1989) adota o conceito de competência comunicativa de Hymes (1972), para marcar posição programaticamente contrária à separação entre competência e desempenho do modelo gerativista. O conhecimento linguístico dos indivíduos inclui não apenas a competência gramatical de gerar e interpretar regras gramaticais, como defendiam os gerativistas adeptos do modelo padrão, mas também a capacidade de utilizá-las de modo eficaz e apropriado em função das convenções da interação verbal. Para Dik (1989), a explicação da competência comunicativa deve basear-se, portanto, em dois sistemas de regras: (i) as que definem a construção das expressões linguísticas, ou componente gramatical; e (ii) as que se centram

nas convenções que regem a comunicação ou interação verbal, ou componente pragmático. Como o primeiro é instrumental em relação ao segundo, devem-se explicar as expressões linguísticas com referência ao sistema de interação verbal.

O postulado de que a sintaxe é empregada para expressar aspectos semânticos e pragmáticos dos enunciados, e que, ao menos nesse caso, seu uso é determinado por esses aspectos, não implica necessariamente o postulado de que a sintaxe também não contenha princípios próprios de organização, que são estritamente independentes. Dik (1983, p.75) parece em princípio admitir isso explicitamente ao responder negativamente à questão se toda a propriedade estrutural deve encontrar uma explicação funcional:

> Como ocorre com todos os instrumentos humanos, as exigências funcionais imputadas a uma língua deixam um pouco de tolerância para especificações alternativas de propriedades não funcionais; desenvolvimentos históricos criam inevitavelmente propriedades rudimentares [...]; e conflitos entre diferentes exigências funcionais podem até criar disfuncionalidade em dadas áreas da organização linguística.[7]

De qualquer modo, o objetivo crucial de Dik foi encontrar um modo de deixar a RS desligada de uma sintaxe autônoma. Segundo Anstey (2004, p.35-6), essa abordagem foi bastante inventiva. Primeiramente, Dik omitiu todas as referências a uma estrutura de constituintes da RS, reduzindo-a aos itens lexicais arranjados numa combinação de relações funcionais, isto é, de relações sintáticas, semânticas e pragmáticas. Em segundo lugar, ele inseriu essa informação oculta em três locais separados, empregando três estratégias: (i) mediante a estratégia de individuação, contornou a

7 Cf. o original: *"As with all human instruments, the functional requirements put on a language leave quite a bit of leeway for alternative specifications of non-functional properties; historical developments unavoidably create rudimentary properties [...]; and conflicts between different functional requirements may even create dysfunctionality in given areas of linguistic organization"*.

necessidade de uma estrutura abstrata na RS, como a de um sintagma verbal, atribuindo aos esquemas de predicado estocados no léxico um número específico de posições argumentais, cada qual com uma função semântica; (ii) mediante a estratégia de semanticização, eliminou a necessidade de um SN criando um mecanismo mediante o qual as relações de dependência interna dos componentes do SN são isomorficamente instanciados na estrutura do termo como relações semânticas; (iii) mediante a estratégia de adiamento, a estrutura da oração foi reificada em *templates* (moldes) de ordem de palavras que não licenciam informação estrutural enquanto a RS não estiver completamente especificada com informação funcional.

Essas estratégias certamente visavam a criar uma oposição sistemática à teoria gerativa, num momento de forte competição, agora superado, entre os enfoques teóricos, pois, de fato, nem mesmo o funcionalismo mais radical afirmaria que a gramática de uma língua não dispõe de pelo menos alguns elementos arbitrários, que fazem parte de um sistema com certa dose de autossuficiência (Croft, 1995, p.494). Nesse caso, o subsistema sintático apresenta princípios próprios e regras independentes, já que estes não se derivam dos princípios e regras dos outros dois subsistemas nem fazem referência a eles. O que defendia a GF, todavia, é uma redução ao mínimo possível desse tipo de fenômeno linguístico. Nesse caso, com relação à autonomia da sintaxe, há uma diferença profunda entre teorias funcionais, como a GF, e teorias formais, como a gerativista.

Apesar da prioridade metodológica do componente pragmático nos estudos funcionalistas, o principal foco de atenção do modelo na GF é a estrutura da oração simples e da oração complexa. Na realidade, o grande esforço de Dik foi dirigido para a explicação da estrutura gramatical com referência à estrutura semântica e pragmática. Ainda que a abordagem dos fenômenos gramaticais recorresse frequentemente a explicações de natureza discursiva, com vistas ao entendimento de que a linguagem é um instrumento de interação social, parece que era a sintaxe em si o objeto principal, não o discurso, já que a unidade máxima estava confinada ao estudo de construções simples (Dik, 1989) e de construções complexas (Dik, 1997).

Nos casos em que a explicação reduz as regras aos mecanismos formais, a GF dá conta apenas da sintaxe, deixando de lado a prioridade das regras pragmáticas e semânticas, um dos postulados mais caros à GF. Com efeito, diz Dik (1989, p.5) que "o correlato psicológico de uma língua natural é a 'competência comunicativa' do ULN [Usuário de uma Língua Natural] no sentido de Hymes (1972): sua capacidade de estabelecer a interação social por meio da linguagem".[8] Apesar disso, Hymes (1979, p.306) lança a seguinte crítica a Dik: "Seria justo dizer que Dik entende ser preferível que a GF seja um modo de analisar a gramática como meio de comunicação, mas deixa a análise dos fins da comunicação e a ligação entre meios e fins para outros ou para um outro momento".[9]

2.4 O discurso na Gramática Funcional nos anos 1990

Desde o início dos anos 1990, uma parte significativa da pesquisa realizada dentro do arcabouço teórico da Gramática Funcional tem sido dirigida para a expansão da GF de uma gramática da sentença para uma gramática do discurso. Connoly (2004) faz um rápido sumário de como a noção de discurso se desenvolveu a partir de 1997 resultando no modelo de Gramática Discursivo-Funcional recentemente proposto.

A primeira abordagem do discurso descrita em Hengeveld (1997), que Connoly menciona, consiste em tratar o discurso mais como um nível adicional (ou conjunto de níveis) acima da oração. Esse enfoque é conhecido como organização em camadas ascenden-

8 Cf. o original: *"The psychological correlate of a natural language is the NLU's 'communicative competence' in the sense of Hymes (1972): his hability to carry on social interaction by means of language".*

9 Cf. o original: *"It might be fair to say that Dik understands FG to be preferable as a way of analysing grammar as communicative means, but leaves the analysis of communicative ends, and the linkage between means and ends, to others, or for another time".*

tes (*upward layering*). Nessa camada adicional, chamada de Retórica, atos de fala são agrupados em Movimentos (*Moves*), que são, por sua vez, organizados num tipo particular de discurso, como a conversação, por exemplo. Um Movimento (*Move*) é realizado por um parágrafo dentro de um monólogo ou por um turno dentro de um diálogo, uma visão de organização do discurso que é, em grande parte, uma retomada do enfoque discursivo de Sinclair e Coulthard (1975), em que atos são agrupados em Movimentos, Movimerntos em trocas, trocas em transações e transações em interações.

A primeira reação a essa proposta de Hengeveld (1997) pode ser encontrada no próprio Dik (1997). Como, todavia, *discurso* pode representar tanto o processo quanto o produto da interação comunicativa, em Dik (1997), como produto, o discurso é concebido como um fragmento coerente de língua falada ou escrita, que contém normalmente mais de uma oração e que pode conter um ou mais constituintes extraoracionais.

Já como processo, o discurso é concebido como interação monodialógica ou dialógica com os participantes sendo rotulados como S para Falante (*Speaker*) e A para Ouvinte (*Addressee*).

Para Dik (1997), qualquer processo discursivo pode ser encaixado dentro de um Evento Discursivo, que é um tipo de fenômeno social, caracterizado por diversos aspectos relevantes, como a identidade dos interlocutores e o relacionamento entre eles; o tempo, o lugar e as circunstâncias do discurso; as convenções sociais predominantes que determinam o uso aceitável da língua. O processo discursivo envolve decisões que afetam a construção do texto em desenvolvimento, o que inclui a escolha do gênero e a seleção do estilo.

Conforme o desenvolvimento do discurso, os participantes criam um quadro mental que habilita a interpretação de cada enunciado novo em relação ao discurso precedente e também em relação ao contexto da informação pragmática dos interlocutores derivada da situação discursiva. Dik (1997) denomina esse grau de compartilhamento entre os interlocutores de Modelo do Discurso. Um Modelo de Discurso tem um caráter dinâmico, por estar em processo de mudança contínua em face das necessidades da interação, e

um caráter parcial, por nunca ser capaz de incluir toda a informação pragmática relevante.

Dik (1997, p.300-4) afirma que cada oração individual tem uma força ilocucionária básica que deve ser vista como uma instrução codificada mediante a qual o falante tenta modificar a informação pragmática do ouvinte. As forças ilocucionárias – declarativa, interrogativa, imperativa e exclamativa – constituem tipos de atos de fala, cuja identificação depende, em grande medida, da diversidade formal com que podem ser codificados na gramática das línguas naturais. Há outros fenômenos, como tempo e tópico discursivo, que permanecem em sequências inteiras de orações ou mesmo na totalidade da superfície do discurso. Por isso, a manutenção da continuidade tópica implica a criação de cadeias de referência anafórica que se estendem para além das fronteiras da sentença.

Dik (1997) postula que, como um evento discursivo é estruturado em duas camadas, a Interpessoal e a Representacional, o discurso tem uma estrutura essencialmente hierárquica, que pode ser, todavia, rompida, por interrupções, encaixamentos e recorrências. A organização hierárquica disposta em camadas pode ser distinguida tanto dentro do Nível Interpessoal quanto dentro do Nível Representacional. No primeiro, a unidade mais baixa é o ato de fala, que veicula um valor ilocucionário determinado. Unidades desse tipo podem combinar-se para formar sequências mais longas de atos de fala. Num diálogo, por exemplo, sequências de um ou mais atos de fala são organizadas em turnos e a sequência completa de turnos pode constituir o discurso.

Do ponto de vista do Nível Representacional, por outro lado, o discurso se divide em episódios e subepisódios, como, por exemplo, capítulos, seções e parágrafos de um livro. Finalmente, essas unidades são compostas por proposições, que constituem o tipo de unidade discursiva de nível mais baixo dentro desse nível de representação. Dik (1997) admite também que relações retóricas como elaboração e contraste, propostas por Mann e Thompson (1987), deveriam ser acomodadas na abordagem discursiva da GF.

Essa proposta de abordagem discursiva não pareceu, no entanto, ainda suficiente para muitos pesquisadores da GF, que passaram a postular uma separação ainda mais nítida entre gramática e discurso do que a representada pelas propostas de Hengeveld (1997) e Dik (1997), sobretudo em defesa da ideia de que o discurso deveria ser tratado dentro de um módulo separado. Assim, o discurso é considerado por Vet (1998), Bolkestein (1992) e Liedtke (1998) como *módulo pragmático* em oposição ao módulo gramatical; por Kroon (1997) como *módulo do discurso* em oposição ao módulo da sentença, e por Steuten (1998) como *unidade comunicativa* em oposição ao conceito de unidade gramatical.

Kroon (1997) argumenta ainda que a unidade mais baixa na hierarquia do discurso, que ela chama de *ato de discurso*, não é necessariamente equivalente ao que Dik (1997) denomina oração (*Clause*), dado que mesmo os chamados constituintes extraoracionais podem atuar como um ato discursivo. Para Vet (1998) e Steuten (1998), a clara diferença de natureza entre as noções de processo e de produto do discurso, mal encoberta pelo conceito formal de oração, é motivo mais que evidente para um tratamento separado no arcabouço descritivo da GF.

Connoly (2004, p.93) afirma que esse desdobramento ocorreu sob o patrocínio da própria GF, já que a ideia de um arcabouço modular parecia, de qualquer modo, já estar prevista por Dik (1997, p.409). Com o falecimento prematuro desse autor, a incumbência de desenvolver uma organização modular ficou destinada a seu discípulo mais brilhante, Kees Hengeveld, que lançou as bases de uma nova arquitetura para a GF, assunto que será discutido a seguir.

2.5 Uma nova arquitetura: a Gramática Discursivo-Funcional

2.5.1 Uma organização modular descendente

Hengeveld (2004a) retorna ao cenário, procurando assimilar aspectos positivos das diferentes propostas. Segundo o autor, há

diversas razões por que a GF deveria expandir-se da sentença para o discurso. Há, em primeiro lugar, muitos fenômenos linguísticos que podem ser explicados somente em termos de unidades maiores que a sentença individual: partículas discursivas, cadeias anafóricas, formas verbais narrativas e muitos outros aspectos da gramática requerem uma análise que leve em conta contextos linguísticos maiores. O exemplo (2-14) mostra um caso de narrativa em que a partícula *ine* no final de uma sentença é sempre retomada na sentença seguinte com a nítida função de dar continuidade ao discurso.

(2-14) *Tidore* (Van Standen, 2000, p.275 apud Hengeveld, 2004c)

turus	*jafa*	*cahi*			*saloi*	*ena=ge*
then	Jafa	carry.on.the.back			basket	3.NH=there
turus	*ena=ge*			*paka*		*ine*
then	3.NH=there		ascend		upwards	
ine		*una*		*oka*		*koi*
upwards		3.SG.M.		pick		banana

'... then Jafa carried the saloi and went upwards. Went upwards he picked the banana'

Há, em segundo lugar, muitas expressões linguísticas menores que a sentença individual, que, todavia, funcionam como enunciados completos e independentes, como frases holofrásticas, exclamações e vocativos, mostrados respectivamente em (2-15a), (2-15b) e (2-15c). Esses fenômenos requerem uma concepção de enunciados como atos discursivos e não como sentenças (Mackenzie, 1998).

(2-15) a *(O que você comprou?) Uma camisa.*
b *Parabéns!*
c *Oh, Maria!*

Hannay e Bolkestein (1998) argumentam que as propostas até então dirigidas à expansão da GF para uma gramática do discurso representavam duas diferentes abordagens. Na primeira abordagem, o nível do discurso é coberto por camadas adicionais hierarquicamente superordenadas. Na segunda abordagem, o nível do discurso é manipulado por um componente separado, ligado ao gramatical mediante uma relação de interface.

Hengeveld (2004a) defende a ideia de que um modelo adequado de gramática do discurso requer a integração dessas duas abordagens, postulando a necessidade de não apenas reconhecer vários níveis de análise, mas também de aplicar uma organização em camadas que se estenda ao nível do discurso. Assim, o modelo de GDF, que o autor apresenta, é tanto modular quanto hierárquico e, além do mais, passa a operar de cima para baixo (*top-down*), significando que as decisões nos níveis e nas camadas superiores determinam e restringem as possibilidades nos níveis e camadas inferiores.

O autor postula ainda que um modelo de gramática é tanto mais efetivo quanto mais sua organização se assemelhar ao processamento linguístico no indivíduo; assim, como a produção começa com as intenções comunicativas procedendo em direção descendente até chegar à articulação da expressão linguística, o modelo de produção gramatical deve refletir esse modo de processamento. É por essa razão que a organização é descendente, partindo da unidade de nível superior, o discurso, até chegar às unidades de nível inferior, unidades de expressão, conforme exposto na Figura 3, que também mostra a organização geral da GDF.

Um parêntese necessário. A elaboração, com suas idas e vindas, da Teoria da Gramática Funcional remonta, como se viu, à década de 1970. É interessante verificar que, em seu estágio atual, os pressupostos teóricos dos seguidores da tradição iniciada por Dik convergem, em grande parte, com os mesmos que os pesquisadores do Projeto de Gramática do Português Falado (doravante PGPF) (cf. Castilho, 2006) constituíram às custas de um processo de convergência parecido com o de montar um carro enquanto ele está em marcha.

O principal ponto de convergência foi justamente a organização da gramática. Em sua primeira versão, a ordem de apresentação dos capítulos mostrava o perfil de abordagem tradicional, com um capítulo inovador. Esse plano, que me furto de mencionar em sua totalidade aqui, apresenta uma Introdução com alguns pressupostos teóricos e uma separação em 4 partes, 1. Fonologia; 2. Classes de palavras e morfologia derivacional; 3. Morfologia flexional; 4. Relações gramaticais; 5. Organização textual-interativa. É ocioso dizer que o capítulo inovador era o quinto.

Essa organização refletia um posicionamento teórico claramente tradicional, com os constituintes mais visíveis do sistema linguístico, as categorias textual-interativas, estrategicamente inseridas no final, e os constituintes menores, construtores, como sintagmas, morfemas e fonemas, inseridos no início da gramática.

O ápice do aludido processo de convergência foi finalmente atingido com uma discussão em torno dos pressupostos teórico-metodológicos, realizada no VII Seminário de 1993. Chegou-se ao consenso de que passariam a representar o trabalho do grupo como um todo os seguintes pressupostos comuns:

1. A concepção de linguagem como uma atividade, uma forma de ação, a verbal, que não pode ser estudada sem se considerar suas principais condições de efetivação.
2. A pressuposição de que, na contingência da efetivação da atividade linguística do falante/ouvinte (i. e., na produção e recepção de textos), temos a manifestação de uma competência comunicativa, caracterizável a partir de regularidades que evidenciam um sistema de desempenho linguístico constituído de vários subsistemas.
3. A pressuposição de que cada um desses subsistemas constituintes do sistema de desempenho linguístico (o discursivo, o semântico, o morfossintático, o fonológico...) é caracterizável em termos de "regularidades" definíveis em função de sua respectiva natureza.
4. A pressuposição de que um dos subsistemas constituintes desse sistema de desempenho linguístico é o subsistema computacional (entendido como uma noção mais ampla que a noção chomskiana

de língua I), definível em termos de regras e/ou princípios envolvidos na organização morfossintática e fonológica dos enunciados que se articulam na elaboração de qualquer texto.
5. A pressuposição de que o texto é o lugar onde é possível identificar as pistas indicadoras das regularidades que caracterizam o referido sistema de desempenho linguístico. (Nascimento, 1993 apud Abaurre; Rodrigues, 2002, p.12-3)

É interessante destacar nesse conjunto a pressuposição de que é o texto o lugar de identificação possível de pistas indicadoras das regularidades que caracterizam a atividade linguística do falante. Essas regularidades não se limitam aos componentes estruturais, tradicionalmente estabelecidos, que compõem o que se denomina "núcleo duro" da linguística, como a fonologia, a morfologia e a sintaxe. Na verdade, o PGPF passou a endossar as descobertas do GT Organização Textual-Interativa, o que significou abandonar as dicotomias tradicionais, como língua-fala e competência-desempenho em nome da linguagem em uso no contexto social de sua produção. Desse novo posicionamento em relação à linguagem resultou um plano totalmente inovador para a Gramática, em que o primeiro volume, já publicado (Jubran; Koch, 2006), trata da Construção do Texto Falado; o segundo volume, também já publicado (Neves; Ilari, 2008), trata das Classes de Palavras e Processos de Construções. O terceiro volume vai tratar da Construção da Sentença; o quarto, da Construção Morfológica da Palavra e, finalmente, o quinto, da Construção fonológica da palavra

Comparando esta com a primeira organização, vemos uma diferença formal que representa um surpreendente salto teórico. Os pesquisadores do PGPF reconheceram na prática a necessidade de formular uma gramática cujo ponto de partida seja seu componente estrutural de maior escopo, o texto e suas relações com a situação discursiva, para chegar a constituintes progressivamente menores da organização linguística. Essa descoberta, já consolidada na prática, só a partir de 2004 recebeu pleno reconhecimento teórico da Teoria da Gramática Funcional, na nova arquitetura projetada por

Henvegeld (2004a), que passou a denominar Gramática Discursivo-Funcional, consolidada em Hengeveld e Mackenzie (2008).

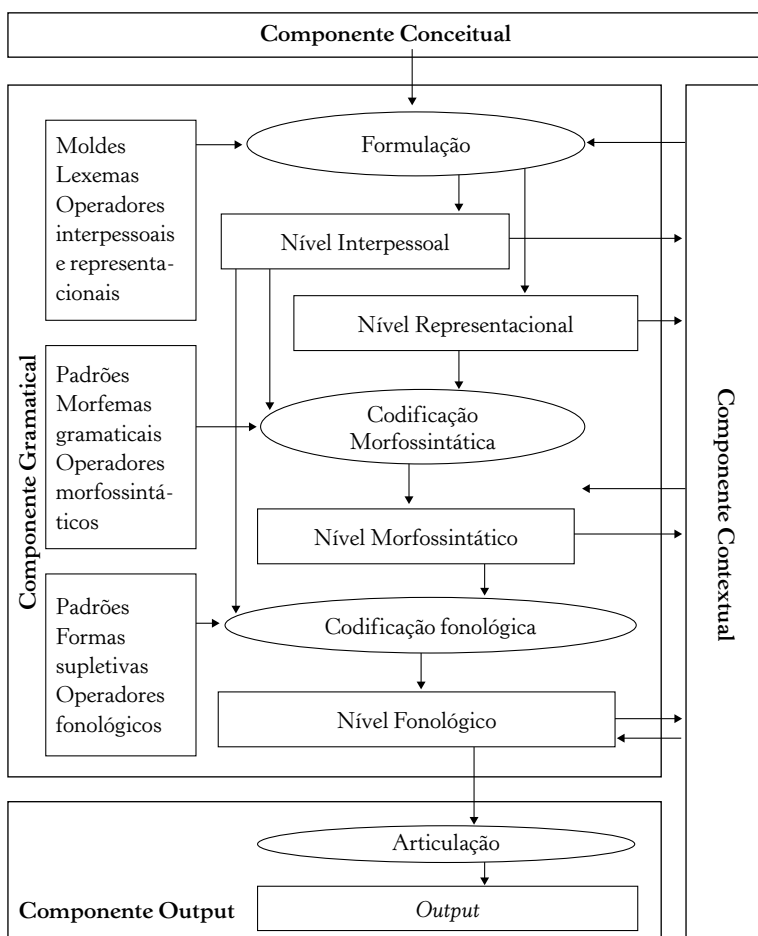
Fechado o parêntese, podemos retomar a arquitetura geral da GDF, num formato nitidamente *top-down*. Hengeveld e Mackenzie (2008) reconhecem a necessidade de uma gramática organizada em quatro níveis, conforme se vê na Figura 4.

(cf. Hengeveld, 2008, p.13)

Figura 4 – Arcabouço geral da GDF

Cada nível – Interpessoal, Representacional, Morfossintático e Fonológico – é concebido como um módulo separado e internamente organizado em camadas. Agindo assim, Hengeveld (2004b) e Hengeveld e Mackenzie (2008) combinam os melhores aspectos das abordagens anteriores.

Outro avanço importante é o reconhecimento de um Componente Contextual e um Componente Conceitual, que contêm elementos essenciais do contexto com o qual os três módulos interagem. O Componente Gramatical é conectado ao Conceitual, ao Contextual e ao componente de expressão (*output*). Assim, representações morfossintáticas e fonológicas representam níveis separados dentro do componente gramatical em vez de constituírem apenas parte do *output* da gramática, como era previsto no modelo *standard* em Dik (1989). Agora, o componente *output* do sistema ficou circunscrito somente à articulação fonética.

Um argumento convincente para justificar a necessidade de um sistema modular com quatro níveis de representação é que é possível fazer referência anafórica, em dois diferentes turnos conversacionais, não apenas a entidades referenciais, como em (2-17), mas também a atos de fala, como em (2-16), à forma de sintagmas nominais, como (2-18), e à própria expressão fonológica, como (2-19).

(2-16) A. *Saia daqui já!*
B. *Não me fale assim!*
Nível Interpessoal (Pragmática)

(2-17) A. *Há muitos sinais de trânsito nesta cidade.*
B. *Eu não havia notado isso.*
Nível Representacional (Semântico)

(2-18) A. *Eu comi "lamb chop" ontem à noite.*
B. *É assim que você diz "costela de carneiro" em inglês?*
Nível Morfossintático (Estrutural)

(2-19) A. *Eu comi uma /pa'eλa / ontem à noite.*
B. *Isso não seria /pa'eya /?*
Nível Fonológico (Fonologia)

(Adaptado de Hengeveld, 2004c)

Todos os níveis têm uma natureza puramente linguística no sentido estrito de que descrevem a linguagem com base nas funções que ela permite exercer, mas somente na medida em que essas funções são codificadas na gramática das línguas particulares; nesses termos, a GDF é, portanto, o componente gramatical de uma teoria mais abrangente da interação verbal.

Como mostra a Figura 3, os fenômenos em que a estrutura linguística é determinada pela interação verbal são capturados como uma interação entre o Componente Gramatical e um Componente Conceitual, um Contextual e um componente terminal (*output*). Dentro do próprio Componente Gramatical, as elipses contêm operações, os quadrados contêm primitivos, isto é, os blocos construtores básicos usados nas operações, e os retângulos contêm os níveis de representação produzidos pelas operações.

O Componente Conceitual pré-linguístico dá igual relevância tanto a uma intenção comunicativa, como, por exemplo, fazer uma advertência, quanto à representação mental correspondente, como, por exemplo, a de um evento causador de perigo. A operação de formulação converte essas representações conceituais em representações pragmáticas (no Nível Interpessoal) e semânticas (no Nível Representacional) que, por sua vez, são traduzidas em representações morfossintáticas e fonológicas mediante a operação de codificação.

Tanto as regras usadas na formulação quanto as usadas na codificação são fenômenos inerentes a línguas particulares, não constituindo, portanto, fenômenos universais. A principal consequência desse princípio é que representações conceituais similares podem aparecer sob a forma de diferentes representações pragmáticas e semânticas em diferentes línguas.

Um exemplo claro dessa diversidade na codificação são os atos de advertência, que, em algumas línguas, recebem uma codificação gramatical própria, constituindo, portanto, um ato de fala distinto, enquanto em outras podem ter a mesma codificação de diretivos. Em português, formular um ato de advertência implica uma extensão da forma gramatical imperativa, usada tanto para diretivos quanto para proibitivos, conforme se vê em (2-20), enquanto o tucano dispõe de uma forma gramatical própria para o imperativo, usado em atos diretivos, e outra para o admoestativo, conforme se vê em (2-21).

(2-20) a *Chegando na casa de sua avó, comporte-se bem, senão...*
b *Faça sua lição já!*

Tucano (Ramirez, 1997, p.145)

(2-21) a *apê-**ya**!*
play-IMP.2PL
'Play!'

b *ape masa tɨro ehâ-gɨ ãyuró wee-**ápa**!*
Another people near arrive-IMPL well do-IMP.ADM
'When you arrive (in the house of) another people, behave yourself' (take care of to do well)

O *output* do Componente Gramatical opera como *input* para a operação de articulação, que, no caso mais comum de expressão, a acústica, contém as regras fonológicas necessárias para um enunciado fonético adequado. Ademais, cada nível dentro do componente gramatical se alimenta do Componente Contextual, que habilita o processo de Referência subsequente a vários tipos de entidade relevantes em cada nível, tão logo sejam eles introduzidos no discurso, conforme demonstrado acima (ver exemplos 2-16 a 2-19). O processo de formulação se apoia nesse componente de modo tal que a disponibilidade de antecedentes e referentes visíveis pode influenciar a composição dos atos discursivos subsequentes.

A Figura 3 especifica também os vários conjuntos de primitivos que alimentam a operação de formulação e a codificação morfossintática e fonológica.

Em primeiro lugar, os esquemas *(frames)* definem as combinações possíveis de elementos nos Níveis Interpessoal e Representacional que são relevantes para a operação de formulação.

Em segundo lugar, o conjunto dos primitivos contém lexemas. Enquanto, na versão anterior, esquemas de predicado são concebidos como elementos diretamente associados ao lexema correspondente, na nova versão, os esquemas são selecionados antes da inserção de lexemas, refletindo as opções disponíveis ao falante para descrever a mesma entidade mediante uma variedade de lexemas com diferentes conotações e/ou denotações.

Em terceiro lugar, esse conjunto de primitivos contém operadores primários, que simbolizam as distinções pragmáticas e semânticas na língua em questão, como, por exemplo, identificabilidade e modalidade, respectivamente.

Todos os níveis da organização na GDF mantêm uma natureza hierárquica e uma organização em camadas, já postulada pela versão anterior da GF, embora a orientação descendente, agora implementada, procure representar mais adequadamente o processamento. Em sua forma máxima, a estrutura geral das camadas dentro dos níveis tem a configuração abstrata contida em (2-22).

(2-22) $(\pi\ v_1\ [\text{head}]\ (v_1)_\Phi]: [\sigma\ (v_1)_\Phi])_\Phi$

(Hengeveld; Mackenzie, 2008, p.14)

Em (2-22), v_1 representa a variável da camada relevante, que é restringida por um possível núcleo complexo e posteriormente especificada por um ou mais modificadores σ, e/ou é posteriormente especificada por um operador π e/ou uma função φ. Modificadores representam estratégias lexicais; já operadores e funções, estratégias gramaticais. A diferença entre estes dois é que somente as funções são categorias relacionais.

2.5.2 Níveis e camadas de organização

O Nível Interpessoal representa os aspectos formais que refletem o papel de uma unidade linguística na interação entre falante e ouvinte. O propósito da interação implica que cada falante emprega uma estratégia, de modo mais ou menos consciente, para atingir seus objetivos comunicativos. Na formulação, o Nível Interpessoal abriga o modo como essa estratégia é realizada em relação tanto aos propósitos do falante quanto ao estado mental corrente do ouvinte. O Nível Interpessoal contém unidades de discurso que são, por sua vez, hierarquicamente organizadas de acordo com a arquitetura global da GDF, como mostra (2-23).

A camada mais alta na hierarquia – o Movimento (*Move*) – constitui uma unidade autônoma de discurso com o efeito perlocucionário de provocar uma reação. Ele descreve um segmento inteiro com as várias camadas inferiores que, por seu lado, contêm componentes desse segmento, uma hierarquia que representa a sequência das ações linguísticas: um Movimento (M) pode consistir em diversos Atos Discursivos temporalmente ordenados. Um Ato Discursivo (A) pode conter diversos Conteúdos Comunicados (C) também cronologicamente ordenados, incluindo, por seu lado, múltiplos Subatos de Atribuição (*Ascription*) (T) e de Referência (*Reference*) (R).

(2-23)

$(\Pi\ M_1: [$ Movimento
$\quad(\Pi\ A_1: [$ Ato Discursivo
$\quad\quad(\Pi\ (F_1: ILL\ (F_1)\ \Sigma\ (F_1))$ Ilocução
$\quad\quad\quad(\Pi\ (P_1: \dots\ (P_1): \Sigma\ (P))_S$ Falante
$\quad\quad\quad(\Pi\ (P_1: \dots\ (P_1): \Sigma\ (P))_A$ Ouvinte
$\quad\quad\quad(\Pi\ (C_1: [$ Conteúdo Comunicado
$\quad\quad\quad\quad(\Pi\ T_1:[\dots]\ (T_1): \Sigma\ (T_1))_\varnothing$ Subato Atributivo
$\quad\quad\quad\quad(\Pi\ T_1:[\dots]\ (R_1): \Sigma\ (R_1))_\varnothing$ Subato Referencial
$\quad\quad\quad]\ (C_1): \Sigma\ (C_1))_\varnothing$ Conteúdo Comunicado
$\quad\quad]\ (A_1): \Sigma\ (A_1))_\varnothing$ Ato Discursivo
$\quad]\ (M_1): \Sigma\ (M_1))$ Movimento

(Adaptado de Hengeveld; Mackenzie, 2008, p.49)

Para dar um exemplo simples, o posicionamento de um Ato Discursivo antes ou depois de outro estrategicamente mais central determina se ele deve ser entendido como uma Orientação ou como um Adendo nos exemplos (2-24a) e (2-24b), respectivamente (cf. Hengeveld; Mackenzie, 2006).

(2-24) a *antigamente* *os cinemas*... *o ambiente era outro* (DID-SP-234)
b *mas chego e já apronto o outro para ir para a escola*... *o menorzinho* (D2-SP-360)

Embora Hengeveld (2004a) preveja a existência de unidades discursivas superiores, o limite para a análise gramatical no Nível Interpessoal é o Movimento, caracterizado como a maior unidade de interação relevante. É possível defini-lo como uma contribuição autônoma para a interação em curso: ele tanto pode requerer uma reação quanto pode ser ele mesmo uma reação. A complexidade do Movimento pode variar muito, do silêncio significativo até um longo fragmento discursivo. Sempre que algum material linguístico estiver envolvido, o Movimento incluirá um ou mais Atos Discursivos; Mackenzie (2004b) sugere que a sentença (2-25) contém três atos discursivos dentro do mesmo Movimento.

(2-25) *Ai meu Deus, 'tá em chamas, meu cabelo!*

(Adaptado de Mackenzie, 2004b, p.183)

Esse enunciado reflete um Movimento em que se manifestam três contornos entonacionais prováveis, cada qual com uma sílaba tônica (*Deus*, *chamas* e *cabelo*, respectivamente). O Movimento pode ser visto como uma sequência de três atos discursivos que refletem naturalmente a sucessão de eventos cognitivos, como choque, percepção e identificação. O autor sugere a representação (2-26) em que a enumeração dos atos corresponde à sequência deles na cognição:

(2-26) $(M_1: [(A_1) (A_2) (A_3)])$

(Mackenzie, 2004b, p.184)

A representação de um Ato Discursivo exibe somente os componentes que foram atualizados pelo Falante, ou seja, minimamente, uma força ilocucionária (F_1) e os participantes do ato de fala $(P_1)_S$ e $(P_2)_A$. Assim, o mesmo enunciado contido em (2-27) pode receber diferentes estruturas subjacentes no Nível Interpessoal, conforme as leituras que sobressaem do contexto da interação, ou seja, ordem ou pedido.

(2-27) *Um café!*

Como, todavia, o Ato Discursivo representa apenas os aspectos que recebem codificação gramatical, cada valor ilocucionário é representado por um operador, e o ato discursivo recebe uma única codificação gramatical em português, que é o ato de fala imperativo, conforme se vê em (2-28).

(2-28) a (M_1: [(Ordem A_1: [IMP $(P_1)_S$ $(P_2)_A$ (C_1: [um café] (C_1))] (A_1))] (M_1))
b (M_1: [(Pedido A_1: [IMP $(P_1)_S$ $(P_2)_A$ (C_1: [um café] (C_1))] (A_1))] (M_1))

(Adaptado de Mackenzie, 1998, p.227)

A terceira camada do esquema ilocucionário é o Conteúdo Comunicado (C), que é constituído por Subatos Referenciais (R) e Subatos Atributivos (A). A introdução da variável C permite agora distinguir a informação comunicada em um ato discursivo (A) e o tipo de entidade a que se faz referência (R) na interação, já que formalizam a função comunicativa do falante. Essas duas entidades se realizam por níveis diferentes e independentes entre si – o Interpessoal e o Representacional. Qualquer tipo de entidade pode representar ora um referente ora um predicado, conforme se observa no exemplo (2-29).

(2-29) a *Ele viu <u>a criança</u>.* (R_1: (x_1) (R_1))
b *Ele viu <u>a criança fazendo um desenho</u>.* (R_1: (e_1) (R_1))

(Garcia Velasco; Hengeveld, 2002, p.110)

O sintagma nominal *a criança* em (2-29a) constitui um Subato Referencial (R) construído a partir da designação de um indivíduo (x), enquanto em (2-29b) é um estado de coisas (e) que contribui para a caracterização de uma entidade referencial (R).

Um exemplo final dá uma visão simplificada da operação do Nível Interpessoal; o Movimento contido em (2-30a) é representado como (2-30b), em que um ato ilocutivo segue um ato interpelativo.

(2-30) a *Ei, você perdeu a carteira!*

b $(M_I$: [

$(A_I: [(F_I: \text{ei } (F_I)) (P_I)_S (P_J)_A] (A_I))$

$(A_J: [(F_J: \text{DECL } (F_J)) (P_I)_S (P_J)_A (C_I: [(R_I) (T_I) (R_J] (C_I))] (A_J))] (M_I))$

(Adaptado de Hengeveld; Mackenzie, 2006)

O Nível Representacional representa os aspectos formais de uma unidade linguística que refletem o papel dela no estabelecimento de uma relação com o mundo real ou imaginado por ela descrito e, desse modo, refere-se mais a uma designação, do que a uma evocação, própria do Nível Interpessoal. O Nível Representacional abriga em categorias formais os aspectos semânticos das unidades linguísticas; em termos formais, as unidades semânticas mediante as quais esse nível atua são organizadas hierarquicamente, de acordo com a arquitetura geral da GDF; pode-se afirmar que as categorias desse nível, que recebem a notação formal contida em (2-31), permanecem praticamente intocáveis desde as versões anteriores da GF.

Deve-se notar, entretanto, que a variável (p), que representa o conteúdo proposicional, constituía a camada superior do Nível Representacional no modelo anterior de oração proposta por Hengeveld (1989), assumida também por Dik (1989), e a notação atribuída era a de um X maiúsculo, conforme mostrado na Figura 2. No entanto, a introdução da variável (C) para conteúdo comunicado no Nível Interpessoal permitiu o deslocamento da variável (p) para o Nível Representacional. Esse nível, construído a partir da inser-

ção de predicados do léxico, representados pela variável (f), que se combinam com entidades referenciais, representadas por (x), permite a expressão de estados de coisas.

(2-31)

$(\pi\ p_1$:	Conteúdo Proposicional
$(\pi\ ep_1$:	Episódio
$(\pi\ e_1$:	Estado de Coisas
$[(\pi\ f_1$: [	Propriedade Configuracional
$(\pi\ (v_1: \blacklozenge\ (v_1): [\sigma\ (v_1)_\Phi])$	Qualquer categoria semântica
. . .	. . .
$]\ (f_1: \sigma\ (f_1)_\Phi])$	Propriedade Configuracional
$(e_1)_\Phi] : (\sigma\ (e_1)_\Phi])$	Estado de Coisas
$(ep_1) : [\sigma\ (ep_1)_\Phi])$	Episódio
$(p_1): [\sigma\ (p_1)_\Phi])$	Conteúdo Proposicional

(Adaptado de Hengeveld; Mackenzie, 2008, p.142)

Na versão mais atualizada da GDF, Mackenzie e Hengeveld (2008) acrescentam as variáveis (l) (t), (m), (r) e (q) para Lugar, Tempo, Modo, Razão e Quantidade, respectivamente, em adição às variáveis (f) e (x), para propriedade e indivíduo, que já existiam antes. As reticências no esquema contido em (2-31) indicam a posição para esses outros tipos de entidades. As diferenças entre as unidades podem ser feitas em termos da categoria ontológica que elas designam. Na medida em que categorias ontológicas têm um reflexo no sistema linguístico, elas adquirem o estatuto de categorias semânticas, cada qual com sua própria variável (Hengeveld; Mackenzie, 2008).

Conteúdos Proposicionais (p), as unidades mais altas do Nível Representacional, representam construtos mentais, como conhecimento, crenças e esperanças. Podem ser factuais, quando eles representam conhecimentos ou crenças racionais acerca do mundo real, ou não factuais, quando envolvem esperanças ou desejos em

relação ao mundo imaginário. Em função de sua natureza, uma característica dos conteúdos proposicionais é que eles podem ser qualificados em termos de atitudes proposicionais (certeza, dúvida e descrença) e/ou em termos de suas fontes ou origens (conhecimento comum compartilhado, evidência sensorial e inferência).

Os Conteúdos Proposicionais contêm Episódios (ep) que constituem, por sua vez, conjuntos de Estados de Coisas; como tais, são tematicamente coerentes por revelarem unidade ou continuidade de Tempo, Locação e Indivíduos. Em diversas línguas a categoria semântica (ep) se manifesta no próprio sistema gramatical como, por exemplo, as que dispõem de ligação Tail-Head. No entanto, mesmo línguas como o português dispõem de uma série de formas verbais não finitas concatenadas como uma narrativa dentro de um discurso maior, que podem funcionar como um (ep).

Estados de Coisas (e), que incluem eventos e estados, se caracterizam por serem localizáveis no tempo e serem avaliáveis em termos de seu estatuto real. Um estado de coisas é caracterizado por uma Propriedade Configuracional (f), que é composicional por natureza e contém uma combinação de categorias semânticas sem relação hierárquica entre si. Essas categorias semânticas de tipos variados podem incluir indivíduos (x), consistindo em objetos concretos que podem ser localizados no espaço e no tempo, e Categorias Lexicais (f), que não têm uma existência independente e podem ser somente avaliadas em termos de sua aplicabilidade a outros tipos de entidades. Outras categorias semânticas, relevantes para a gramática das línguas individuais, que podem participar da constituição de uma Propriedade Configuracional, são Locação (l), Tempo (t), Modo (m), Razão (r) e Quantidade (q).

A descrição de uma entidade pode tomar a forma de um núcleo lexical, como em (2-32), ou a forma de um núcleo complexo, que consiste em uma combinação de outras categorias semânticas, como em (2-33).

(2-32) $(1\ x_i: (f_i: menino_N\ (f_i)_\varphi)\ (x_i)_\varphi)$
'um menino'

(2-33) (Pass e_i: [(f_i: ler_V (f_i)) (1 x_i: $menino_N$ $(x_i))_{Ag}$ (1 x_j: $livro_N$ $(x_j))_{Pac}$] (e_i))
'O menino lê o livro.'

Em (2-32), um nome, em si mesmo designando uma propriedade (f_i), fornece uma descrição lexical simples de um indivíduo (x_i); já em (2-33), uma combinação de categorias semânticas, introduzida entre colchetes, fornece uma descrição composicional de um Estado de Coisas. Assim, as combinações de categorias semânticas, tanto em termos quantitativos quanto em termos qualitativos, são especificadas com base em esquemas representacionais, que formam parte do conjunto de primitivos que alimenta o processo de formulação.

O Nível Morfossintático dá conta dos aspectos estruturais de uma unidade linguística seja em relação à estrutura de sentenças, orações e sintagmas, seja em relação à estrutura interna das palavras. Juntamente com o Nível Fonológico, ele dá conta da codificação das distinções decorrentes da atuação dos Níveis Interpessoal e Representacional.

Em vista dessa operação codificadora, o Nível Morfossintático acaba sendo funcionalmente motivado: princípios de ordenação de constituintes são motivados por relações de iconidade, de integridade de domínio e por preservação de relações de escopo. Esses aspectos não impedem, todavia, que o Nível Morfossintático tenha, ao mesmo tempo, seus próprios princípios de organização, como ocorre, por exemplo, com a imposição arbitrária de um padrão de ordenação não funcionalmente motivado. Como a GDF não estabelece distinção entre sintaxe e morfologia, os princípios aplicados à formação de palavras são os mesmos que se aplicam à formação de sintagmas e orações.

O conjunto de primitivos usados na codificação morfossintática fornece os padrões estruturais (*templates*) apropriados com base nos quais se estrutura o Nível Morfossintático. Como os outros níveis, o Morfossintático é hierarquicamente organizado e, na notação formal, usam-se estruturas em colchetes e rótulos categoriais para capturar as propriedades formais relevantes das unidades linguísticas, conforme a configuração formal contida em (2-34).

(2-34)

(Le_1:	Expressão Linguística
(Cl_1:	Oração
(Xp_1:	Sintagma
(Xw_1:	Palavra
(Xs_1)	Raiz
(Aff_1)	Afixo
(Xw_1))	Palavra
(Xp_1))	Sintagma
(Cl_1))	Oração
(Le_1))	Expressão Linguística

(Adaptado de Hengeveld; Mackenzie, 2008, p.291)

Uma Expressão Linguística é qualquer conjunto de pelo menos uma unidade morfossintática. As unidades combinatórias no interior da Expressão Linguística podem ser orações, sintagmas ou palavras. Dada a orientação funcional da GDF, há uma clara expectativa de que propriedades de ordenação possam ser explicadas com base no uso e no significado das unidades linguísticas. Há, contudo, muitas outras propriedades de ordenação governadas por princípios independentes, aspecto que garante a validade de um nível morfossintático independente, em vez de ser apenas o *output* da gramática, como aparece na GF (Dik 1989). Outra motivação para a presença desse nível dentro da gramática é o fato de ser possível fazer referência anafórica a unidades morfossintáticas conforme já explicitado acima em relação ao exemplo (2-18).

O Nível Fonológico admite *input* do Nível Interpessoal até o Morfossintático e fornece *input* para os processos extragramaticais de articulação, mas a GDF se concentra apenas nos aspectos da fonologia que refletem o funcionamento da linguagem no processo de comunicação.

Como os outros níveis, as representações fonológicas têm também uma natureza hierárquica e também a esse nível se aplica o postulado da GDF de que nem todas as camadas estão ativas em

todo Enunciado ou é de fato relevante para todos os sistemas linguísticos. A organização máxima em camadas que a GDF admite é a que se vê em (2-35).

(2-35)

$((\pi\ U_1: [$	Enunciado
$(\pi\ IP_1: [$	Sintagma Intonacional
$(\pi\ PP_1: [$	Sintagma Fonológico
$(\pi\ PW_1:[$	Palavra Fonológica
$(\pi\ F_1: [$	Pé
$(\pi\ S_1))^N$	Sílaba
$]\ (F_1))$	Pé
$]\ (PW_1))$	Palavra Fonológica
$]\ (PP_1))$	Sintagma Fonológico
$]\ (IP_1))$	Sintagma Intonacional
$]\ (U_1))$	Enunciado

(Adaptado de Hengeveld; Mackenzie, 2008, p.428)

A introdução de quatro níveis independentes de análise permite simplificar certos aspectos da representação de conformidade com o postulado funcional do modelo de não admitir transformações. A tradição recente da GF ativou certo grau de polêmica em torno da representação de expressões linguísticas, que, apesar de não constituírem uma unidade sintática oracional, apresentam independência formal na interação verbal por desenvolverem um ato de fala. A tendência de Dik (1989) é introduzir na representação todos os estratos hierárquicos da oração, mesmo que o falante seja incapaz de reconstruir cada um deles a partir do contexto.

Entretanto, em consonância com a orientação descendente da GDF, Hengeveld (2004a) entende que é a própria intenção comunicativa o primeiro estágio na criação de uma expressão linguística; por conseguinte, a expressão de frustração numa dada situação comunicativa pode manifestar-se diretamente da seleção de um item lexical, como, por exemplo, a interjeição "droga!", que é, por

si mesma, destituída de conteúdo semântico; por essa razão, não há motivo para que o Nível Representacional contenha algum tipo de informação na RS dessa expressão, como se vê em (2-36)

(2-36) *Droga!*

A_1: [EXPR $(P_1)_S$ $(P2)_A$	(C_1)	] (A_1))

	(Lex_1)	

(Adaptado de Hengeveld, 2004a, p.11)

Considere-se, agora, um caso em que o falante se apoia em categorias do Nível Representacional para transmitir sua intenção comunicativa. O conteúdo comunicado (C) é a descrição de uma entidade de terceira ordem (p) expressa numa oração (C_1):

(2-37) *O Ibirapuera é o melhor lugar para se fazer caminhadas.*

A_1: [DECL $(P_1)_S$ $(P2)_A$ (C_1: [	(C_1)	] A_1))
	(p_1)	
	(Cl_1)	

(Adaptado de Hengeveld, 2004a, p.11)

Por outro lado, o mesmo conteúdo proposicional (p) expresso numa oração (Cl_1) pode ocorrer como o veículo para o falante executar um ato referencial (R).

(2-38) *Quero saber se o Ibirapuera é o melhor lugar para se fazer caminhada.*

A_1: [DECL $(P_1)_{Sp}$ $(P2)_{Addr}$ (C_1: [	(R_1)	] (C_1))] A_1))
	(p_1)	
	(Cl_1)	

(Adaptado de Hengeveld, 2004a, p.11)

A verdade é que a orientação descendente do modelo privilegia o princípio de adequação psicológica, defendido desde o início pela GF; outro aspecto também relevante é o de que essa orientação, vinculada a certa simplicidade notacional, respeita também o princípio de economia, dado que as representações incluem apenas a informação necessária para a produção da expressão. Mesmo assim, a proposta original de Hengeveld (2004a) recebeu muitas críticas dos pesquisadores que trabalham com o mesmo arcabouço teórico.

Hengeveld (2004b) lembra que, desde a introdução da estrutura em camadas, questiona-se se ela não deveria ser interpretada como uma representação do próprio processo comunicativo (Bolkestein, 1992; Harder, 1996). Muitos pesquisadores levantaram o mesmo problema em relação às soluções propostas pela GDF. Hengeveld (2004b) rebate, assinalando que a GDF é um modelo que representa fatos linguísticos e, como tal, é apresentada como o componente gramatical de uma teoria mais abrangente da interação verbal; por isso, o componente conceitual e o contextual não fazem parte do componente gramatical em si, mesmo que muitos fenômenos gramaticais possam ser estudados mais adequadamente se se assumir a existência desses componentes (cf. Hengeveld, 2004b, p.366).

Outro aspecto que vale a pena ressaltar é que alguns autores, como Bakker e Siewierska (2004), Harder (2004) e Mackenzie (2004b), enfatizam a importância de uma implementação dinâmica do modelo, o que significaria ser a gramática o reflexo direto do processo de produção. Uma vez mais, rebate Hengeveld (2004b) que postular que a gramática é parte do processo de interação verbal não implica entendê-la como um modelo do falante, embora seja lícito assumir que a efetividade do modelo é diretamente proporcional ao grau de semelhança que tiver com o processo de produção. Entende Hengeveld que, para a implementação da GDF, esse postulado significa que os vários níveis operam simultaneamente, através de uma reduzida diferença temporal entre o nível mais alto e o imediatamente inferior.

Outro ponto que cabe assinalar é que, para Hengeveld (2004b), a implementação dinâmica se depararia com um importante aspec-

to envolvendo os princípios básicos da GF. Conforme postulam Bakker e Siewierska (2004), a ordem dos constituintes no Nível Morfossintático produzida dinamicamente acaba tendo que ser linear, em oposição ao princípio postulado por Dik (1989) de que as representações subjacentes devem ser não ordenadas. A consequência disso está no fato de que, no tipo de implementação dinâmica postulado por esses autores, somente restam duas alternativas: ou as representações semânticas têm que ser especificadas antes da expressão, ou também elas devem ser linearmente ordenadas. A opção de Bakker e Siewierska (2004) pela segunda alternativa não apenas difere drasticamente da tradição corrente na GF, mas parece sugerir que diferenças na organização estrutural refletem diferenças na organização semântica. Segundo Hengeveld (2004b), a existência de pronomes catafóricos e reflexivos que precedem seus referentes são fenômenos que contradizem esse tipo de implementação dinâmica.

Quanto às funções, a hierarquia funcional de determinação pragmática > semântica > sintaxe instancia-se na ordenação dos níveis Interpessoal, Representacional e Morfossintático, respectivamente, que fornecem as posições estruturais relevantes para funções pragmáticas, semânticas e sintáticas. A atribuição de funções sintáticas pode ser encarada, portanto, como o resultado de um processo que leva em conta tanto fatores pragmáticos, situados no Nível Interpessoal, quanto fatores semânticos, situados no Nível Representacional. Essa proposta tem consequências relevantes para a concepção de funções sintáticas na GF, considerando que elas devem ser vistas como noções puramente gramaticais e não semânticas. A noção de perspectivização prevista nas funções sintáticas de Sujeito e Objeto é, portanto, o produto de propriedades pragmáticas e semânticas mais básicas, que acionam o aparecimento dessas funções sintáticas, em vez de fazerem parte do significado, como se via antes na GF de Dik (1989).

Como conclusão aqui, é possível afirmar que a GDF sucede a GF na linha do trono (Dik 1989; 1997), mas não a substitui completamente. Por um lado, expande o escopo da GF por levar rigorosamente em conta os princípios de adequação pragmática e psi-

cológica, ao propor que a instrumentalidade do sistema linguístico em criar e manter relações comunicativas (adequação pragmática) é real, mas obedece a restrições cognitivas gerais sobre a produção e a interpretação do discurso (adequação psicológica). Além disso, a nova arquitetura retém a força original da GF mediante a combinação de neutralidade tipológica e rigor formal. É supérfluo dizer que adotar, como ponto de partida, o princípio da intenção comunicativa do falante em influenciar o interlocutor por meio do discurso é um objetivo que impõe à teoria condições explanatórias de natureza extralinguística.

De acordo com Hengeveld e Mackenzie (2006), a tipologia linguística, entendida como o estudo dos princípios subjacentes à variação formal entre as línguas do mundo, é uma fonte essencial de inspiração para o novo modelo. Isso significa que a GDF mantém o objetivo da GF de desenvolver um arcabouço para a descrição sistemática de todas as línguas humanas possíveis, em atenção ao princípio de adequação tipológica proposto por Dik (1989).

Além disso, entendem os autores que a GDF representa um avanço para os estudos tipológicos. Ao postular distinções formais derivadas da atuação de diferentes níveis, a GDF força, de certa maneira, o tipologista a ultrapassar o nível estrutural da descrição tipológica, geralmente restrita a considerações de ordem sintática e morfológica, para avançar na direção do estudo sistemático de uma tipologia semântica e pragmática das línguas naturais.

O aparecimento de uma notação diferente não deve significar também que mudaram os compromissos cruciais da GF. Pelo contrário, a GDF, apesar de toda a sofisticação notacional que a sustenta, é ainda basicamente uma teoria não transformacional de gramática, mediante a qual expressões linguísticas são geradas por regras de expressão operando sobre representações subjacentes; estas, por sua vez, identificam relações funcionais entre itens lexicais embutidos em esquemas diversos, hierarquicamente organizados um dentro do outro.

Essa continuidade entre o modelo de Hengeveld (2004a) e as fases anteriores da GF não deve obscurecer as descontinuidades

relevantes. Anstey (2004, p.45) entende que há duas características definidoras na nova versão do modelo: a mudança da centralidade no predicado para a centralidade na pragmática e a divisão da representação subjacente da GF tradicional nos Níveis Interpessoal, Representacional, Morfossintático e Fonológico.

Encerrada essa apresentação do suporte teórico, fica para o próximo capítulo discutir o modo como esse aparato ajuda a definir a nominalização. Como a GF não é unânime quanto a esse aspecto, ele requererá a definição de um posicionamento teórico.

3
UM ENFOQUE FUNCIONAL DA NOMINALIZAÇÃO

Introdução

Conforme especificado na Introdução, o principal objetivo deste trabalho é sustentar a hipótese de uma continuidade categorial entre classes de palavras, tomando como parâmetro a diferença entre o nome e o verbo, mediada pela nominalização, categoria aqui postulada como constituindo um termo complexo entre os dois outros opostos, tomados como membros prototípicos.

Nesse caso, a sustentação dessa hipótese passa necessariamente pela comprovação de que as nominalizações, cuja função categorial é referir-se a entidades de segunda ordem ou de ordem superior, dispõem de estrutura argumental, mesmo nos casos em que os argumentos não são superficialmente especificados.

É a descrição das nominalizações em uso nas situações reais de fala o aspecto crucial para dar sustentação a essa hipótese de continuidade funcional, tarefa que desenvolvo no Capítulo 4. Nesse caso, a finalidade do presente capítulo é a de fornecer um embasamento teórico para a nominalização no âmbito do modelo funcional adotado, cujos parâmetros epistemológicos já foram apresentados no capítulo anterior.

O conteúdo se organiza do seguinte modo: na Seção 3.1, discuto a concepção dikiana de construção encaixada e, por conseguinte,

a inserção da nominalização na classe de predicados encaixados. A Seção 3.2 destina-se à apresentação das duas alternativas desenvolvidas na GF para o tratamento da estrutura argumental da nominalização: a hipótese de preservação de valência e a hipótese de redução de valência. Finalizo o capítulo com a Seção 3.3, que é dirigida ao desenvolvimento do modo como a nominalização passou a ser tratada no arcabouço da GDF.

3.1 A nominalização como construção encaixada

Nominalizações e outros tipos de predicados constituem construções encaixadas. Embora haja, de um ponto de vista tipológico, uma grande variedade de tipos, é possível discernir padrões recorrentes de construções encaixadas. Dik (1997) fornece uma taxionomia com validade transistêmica para as construções encaixadas, que podem ser vistas na Figura 5.

Embora o único traço que as construções infinitivas e participiais (gerúndios e particípios) têm em comum com as nominalizações seja a possibilidade de constituir o núcleo da construção encaixada, há outro traço relevante, compartilhado por todos os tipos de construções encaixadas não finitas, que é a possibilidade de realização como predicação fechada ou aberta, conforme a natureza especificada ou não especificada da estrutura argumental de que dispõem.

No entanto, Dik (1997) atribui esse traço somente a construções infinitivas e participiais, deixando de fora as nominalizações, embora as inclua entre as predicações encaixadas que têm propriedades nominais. Mais adiante, defendo a ideia de que a noção de predicação aberta é crucial para se entender a ausência de argumentos nas nominalizações; portanto, essa propriedade deve ser atribuída às nominalizações também. Na parte II, que trata da investigação empírica, fornecerei evidências favoráveis a essa posição, cuja consequência teórica mais relevante será uma nova proposta para a taxonomia das construções encaixadas contida na Figura 5.

(cf. Dik, 1997, p.142)

Figura 5 – Tipos de construções encaixadas

Por nominalização, Dik (1997, p.157-8) identifica um tipo de construção encaixada que tem uma ou mais propriedades em comum com um termo nominal primário. Como uma construção encaixada, a nominalização preenche a posição de algum termo em uma construção hierarquicamente superior, que atua como predicação matriz. Considere os exemplos em (3-13a) e (3-13b)

(3-13) a *Maria lamentou que João tenha demitido Pedro.*
b *Maria lamentou a <u>demissão</u>* ***de Pedro*** *por João.*

---

1 Construções participiais incluem tanto os chamados particípios passados quanto os chamados particípios presentes, que, como é sabido, são praticamente ausentes do paradigma verbal do português, que emprega o gerúndio, para desempenhar funções correlatas. Nesse caso, esse tipo de construção encaixada deve ser visto como abarcando particípios e gerúndios, que têm uma função similar no português, no que tange ao caráter aberto ou fechado do predicado envolvido.

No caso de nominalizações, como (3-13b), é o próprio predicado que adquire certas propriedades nominais; com efeito, vê-se que o constituinte em função de Paciente de (3-13a) aparece, na nominalização de (3-13b), sob a forma de um sintagma de possuidor, que é especialmente adequada para expressar relações no interior de termos nominais. O fato de *Pedro* aparecer em (3-13b) como sintagma de possuidor é interpretado por Dik como uma propriedade nominal, e o SN *a demissão de Pedro* é descrito como uma nominalização (Dik, 1997, p.157-8).

Dik distingue também nominalizações nucleadas e não nucleadas. No tipo nucleado de nominalização, a construção encaixada é adicionada ao núcleo nominal, que tem um significado geral, como "fato" ,"coisa", ou "circunstância", como se observa em (3-14a). Já na nominalização não nucleada, é o predicado em si mesmo que assume certas propriedades nominais, inclusive posições argumentais numa estrutura complexa, como se observa em (3-14b).

(3-14) a *Maria considerou o fato de que poderia ser derrotada nas eleições.*
b *Maria considerou sua derrota nas eleições.*

Conforme já mencionado anteriormente, para a GF, podem-se preencher com termos as posições de argumentos e de satélites na RS da oração. Termos permitem construir a representação mental de uma entidade qualquer, ou identificar/retomar uma entidade já estabelecida em seu mundo mental (Dik, 1997). O tipo primário de termo é representado em geral por um nominal que se refere a alguma entidade de primeira ordem, e o modelo de expressão prototípico para esse termo inclui constituintes como determinante, quantificador, possuidor, adjetivo, nome. O exemplo contido em (3-15a) mostra um SN e respectiva formalização para determinante e modificador do núcleo nominal, contida em (3-15b).[2]

---

2 Na RS contida em (3-15b), o determinante é um operador e, como tal, aparece designado por **d** (representando o traço semântico "definido"); as regras de expressão são responsáveis pela inserção do artigo no enunciado final, que constitui o *output* da gramática (cf. Dik, 1989).

(3-15) a *o deputado corrupto*
b $d1x_i$: deputado [N] $(x_i)_\emptyset$: corrupto [A] $(x_i)_\emptyset$

Em referência a entidades de ordem superior, os termos podem ser representados por nomes simples ou por termos complexos operando como restritores, que constituem construções contendo uma predicação ou uma proposição. Os exemplos contidos em (3-16a-b) ilustram termos complexos com o mesmo núcleo que fazem referência a entidades de duas diferentes ordens superiores:

(3-16) a *Chaves testemunhou* <u>*a derrota de Bush.*</u>
b *Chaves acreditou* <u>*na derrota de Bush.*</u>

É a natureza semântica do predicado que determina o nível a que cada termo se refere. Desse modo, em (3-16a), a predicação [Chaves testemunhou] – interpretada como designando um estado de coisas – faz referência a um evento, uma entidade de segunda ordem, e, portanto, atua no nível da predicação (ver Figura 2). Já (3-16b) tem um estatuto semântico diferente: as coisas em que pessoas acreditam não podem ser classificadas como um estado de coisas, mas como um conteúdo proposicional ou fato possível que constitui uma entidade de terceira ordem (Dik, 1989; 1997).

Comparando-se os nomes destacados de (3-16) com o de (3-15), observa-se uma diferença fundamental: apenas os de (3-16) são de ordem superior, no mínimo de segunda ordem, enquanto o de (3-15) constitui um tipo primário de termo, cuja função prototípica é referir-se a entidades de primeira ordem. Todos os exemplos de (3-16), que são casos de nominalizações, constituem um tipo secundário de termo, cujo papel primordial é atuar como construção encaixada.

Dik (1985) procura dar uma explicação plausível para os reflexos entre forma e conteúdo na formação de predicados derivados, postulando dois princípios, que ele denomina Princípio de Ajuste Formal (doravante PAF) e Princípio de Ajuste Semântico (doravante PAS). O poder explanatório desses dois princípios se assenta no fato de permitirem fazer predições corretas não apenas

em relação a manifestações de construções causativas em diferentes línguas, mas também em relação a outros tipos de construção, tais como as nominalizações, processo diretamente vinculado aos objetivos deste trabalho.

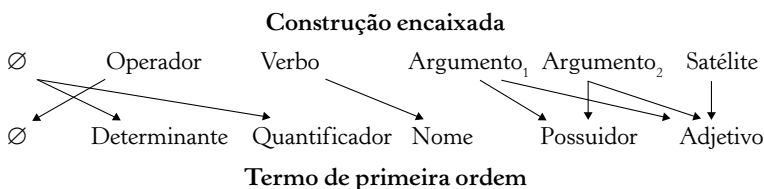
Com base nos princípios acima expostos, Dik afirma que é o PAF que explica como predicações basicamente verbais adquirem propriedades nominais, isto é, descategorização e recategorização: uma predicação encaixada atua como um termo de uma predicação mais alta, e, por seu lado, termos se definem como expressões nominais. A Figura 6 representa os ajustes formais.

(Dik, 1997, p.158)

Figura 6 – Ajustes formais entre verbos e nomes

Segundo Dik (1985), o PAF e o PAS governam as expressões formais e as propriedades semânticas das construções derivadas. Por um lado, o PAF prediz que construções derivadas devem ajustar sua expressão formal ao Modelo de Expressão Prototípico (doravante MEP), fornecido por construções não derivadas. O PAS, por outro lado, tende a se ajustar também às propriedades semânticas do MEP à medida que uma construção derivada se submete à pressão do PAF. Esse modelo de ajuste se baseia nas definições contidas em (3-17):

(3-17) a predicados de zero-lugar não têm MEP.

b para predicados de um lugar, o MEP é o modelo de expressão para predicados de um lugar com um primeiro argumento específico: *João* *caminha*/ *a* *caminhada* *de João.*

c para predicados de dois lugares, o MEP é o modelo de expressão para predicados de dois lugares com um argumento Agente e um argumento Meta: *João comprou o livro/ a compra do livro por João.*

d para predicados de três lugares, o MEP é o modelo de expressão para predicados de três lugares com argumentos Agente, Meta e Recipiente: *João ofereceu o livro a Pedro/ a oferta do livro por João a Pedro.*

e predicados de quatro lugares: não têm MEP.[3]

f para termos, o MEP é o modelo de expressão para um termo com um nominal não derivado como núcleo, possivelmente modificado por adjetivos atributivos e sintagmas de possuidor, e determinado por um ou mais operadores de termo, conforme os exemplos: *a casa velha do meu amigo/a lenta destruição de Cartago*.

Com base nos princípios acima, Dik afirma que o PAF explica o modo como predicações basicamente verbais adquirem propriedades nominais: uma predicação encaixada atua como um termo de uma predicação mais alta e, por seu lado, termos se definem como expressões nominais. Conforme esquematizado na Figura 6, são os seguintes os ajustes mais comuns da predicação verbal encaixada à expressão do termo nominal:

(i) um operador de predicado, como o morfema de número, pessoa e modo-temporal, passa a zero e, inversamente, zero se torna operador de termo, conforme se observa em (3-18a-b): em (3-18a), o operador de predicado (tempo pretérito, por exemplo) passa a zero e (3-18b) recebe o acréscimo de um operador de termo, que é o determinante *a*:

3 De maneira geral, a valência do nome derivado corresponde à do predicado *input*. Dik (1985) considera que não há MEP para predicados tetravalentes justamente porque não há SN (nominal não derivado) de quatro lugares que possa servir de modelo de expressão a esses predicados.

(3-18) a *O presidente demitiu o ministro tardiamente.*
b *A demissão tardia do ministro pelo (*do) presidente.*[4]

(ii) um predicado verbal torna-se núcleo nominal: o predicado verbal de (3-18a), com a RS
demitir$_v$ *(presidente)*$_{Agente}$ *(ministro)*$_{Meta}$, , passa a predicado nominal, em (3-18b) com a RS
demissão$_N$ *(presidente)*$_{Agente}$ *(ministro)*$_{Meta}$.;

(iii) o primeiro e o segundo argumentos podem tanto assumir a forma de um sintagma de possuidor quanto a de um adjetivo atributivo: em (3-19b) o argumento de (3-19a) é expresso sob a forma de sintagma de possuidor e, em (3-19c), sob a forma de um adjetivo;

(3-19) a *Elegeu-se o presidente.*
b *A eleição do presidente.*
c *A eleição presidencial.*

(iv) um satélite adverbial só pode assumir a forma de um adjetivo atributivo, o que aparece claramente demonstrado em (3-18b): o advérbio de tempo *tardiamente* de (3-18a) assume a forma de expressão de modificador típico.

O aspecto formal mais saliente, e, por isso mesmo, o que Dik discute mais detalhadamente, concerne ao ajuste de argumentos à expressão de sintagmas de possuidor. Para Dik, é a forma de expressão de possuidor (genitivos, sintagmas-de e pronomes possessivos) que constitui o modelo mais comum para a expressão de argumentos nas nominalizações.

---

4 Em português, a expressão nominalizada não faz jus ao princípio de ajustes formais mencionado por Dik (1997), já que dificilmente a entidade agentiva recebe o mesmo tratamento formal de sintagma-de que a entidade afetada, ou Meta como em *a demissão do ministro pelo (*do) presidente.* Observe, no entanto, que línguas com genitivo como o inglês permitem agentes no papel de genitivo: *the president's dismissal of the War Ministry.*

Os argumentos ajustados são tipicamente o primeiro e o segundo argumento em uma predicação transitiva, que exercem função semântica de Agente e Meta, respectivamente, que competem entre si para ocupar a posição de possuidor. Enquanto, nos predicados monovalentes, o ajuste é automático, nos predicados bivalentes, os argumentos disputam a única posição disponível de possuidor.

Há uma tendência inequívoca, detectada por Du Bois (1987) no sacapulteco, Asby e Bentivoglio (1993), no francês e no espanhol, e por Dutra (1987), Bentivoglio e Braga (1988) e Pezatti (1992) no português, para que as orações contenham um único argumento lexical, geralmente na posição de Meta, fenômeno expresso por uma regra de estrutura preferida do tipo "evite mais de um argumento lexical na oração" (Du Bois, 1987).

Dik (1985) observa um comportamento semelhante das predicações nominais derivadas numa amostragem do holandês, em que ocorre um único exemplo de predicado transitivo com a expressão tanto de Agente quanto de Meta. Isso implica que, similarmente à regra de um argumento lexical na predicação básica, não derivada, a expressão da regra de estrutura preferida nas nominalizações pode resumir-se a "expresse preferencialmente o argumento como um sintagma de possuidor" (Dik, 1985, p.25). O termo "preferencialmente" foi usado em função do fato de o português necessitar inserir dois sintagmas-de para a expressão do possuidor em contraste com o inglês, por exemplo, que dispõe de argumento-de e genitivo para cada posição de possuidor.

As regras para a expressão de argumentos da nominalização em português recebem a seguinte formulação, adaptada de Dik (1985, p.24):

1. Se há um primeiro argumento compatível com as condições exigidas por possuidores pré-nominais, então: (a) se expressa o primeiro argumento como um possuidor pré-nominal (regra opcional).[5]

5 No caso do português, isso ocorre por meio do emprego de pronome possessivo, possibilidade estendida para genitivos em outras línguas, como o inglês e o holandês, por exemplo.

2. Se há um termo na função de Meta, então, (b) se expressa o argumento Meta como um sintagma-de (regra obrigatória).
3. Se, havendo um primeiro argumento, não se aplicam nem (1) nem (2), então, (c) se expressa o primeiro argumento num sintagma-de (regra opcional, mas preferida).
4. Expressam-se termos não especificados em a-c acima de acordo com sua própria função semântica.[6]

As regras não se aplicam da mesma forma, já que a opção entre (a) e (c) equivale, na 3ª pessoa, a uma opção entre um possessivo e um sintagma-de. A condição anafórica do possessivo implica que haja uma menção anterior do termo correferente, diferentemente do sintagma-de, que implica, geralmente, menção pela primeira vez do termo referido.

Observe-se, entretanto, o efeito dessas regras, no português, em predicações encaixadas de um lugar, como em (3-20):

(3-20) $(x_i$: [falir (o sistema bancário)$_{\text{Processado}}$] $(x_i))$

Satisfaz-se a condição 1, já que se pode expressar *O sistema bancário* como um possuidor prenominal; selecionada a opção (a), o resultado é, em português, *sua falência* (com as condições textuais naturalmente impostas a um termo anafórico como *sua*). Não selecionada a opção (a) e satisfeita a condição 2, aplica-se, então, a opção (c). O resultado é, em português, *a falência do sistema bancário* (considerando ser *o sistema bancário* um termo mencionado pela primeira vez).

Considere-se agora um predicado de dois lugares, como em (3-21):

(3-21) $(x_i$: [corromper$_V$ (o deputado)$_{\text{Agente}}$ (o funcionário)$_{\text{Paciente}}$] $(x_i))$

6 É conveniente lembrar verbos como *conviver*, que regem preposições específicas. A nominalização correspondente herda a preposição introdutora do argumento oblíquo: *a convivência de José <u>com Maria</u>*. Nesse caso, obviamente os argumentos não competem entre si pela posição de possuidor, já que cada um recebe sua própria forma de expressão.

Novamente, satisfaz-se a condição 1. Aplicada a opção (a) e a regra obrigatória (b), o resultado é *sua corrupção do funcionário*. Caso não se apliquem as regras (a) e (c) à nominalização, o resultado seria *a corrupção do funcionário pelo deputado*.

Esse sistema apresenta, assim, as seguintes características:

1. Há uma forte preferência para que, pelo menos, um argumento central assuma a expressão de possuidor.
2. O argumento na função de Paciente, se houver, deve assumir a expressão de possuidor, em posição pós-nominal.
3. Embora esse conjunto de regras possa produzir uma construção com dois argumentos centrais na expressão de possuidor, um deles ocupa necessariamente a posição pós-nominal e o outro, obrigatoriamente o primeiro argumento, ocupa a posição pré-nominal (Dik, 1985, p.26).[7]

Para Dik (1997), é incomum a ocorrência de dois argumentos centrais da nominalização como expressões possessivas do mesmo tipo. Considerem-se os exemplos do inglês, contidos em (3-22a-c).

(3-22) a *The signing of the cheque of John.*
b *John's signing of the cheque.*
c *The signing of the cheque by John.*

A construção em (3-22a) não é licenciada pela gramática porque se usam duas expressões possessivas do mesmo tipo sintático. Para expressar o conteúdo na formulação de (3-22a), teriam de ser usados (3-22b) ou (3-22c).

7 Cf. o original: *"1. there is a strong preference for at least one central argument to take possessor expression. 2. if there is a Goal, the Goal must take possessor expression; this can only be done in postnominal possessor position. 3. the constellation of rules may yield a construction with two central arguments in possessor expression, but this is only possible when one of them is a prenominal possessor and the other a postnominal possessor, and when the prenominal possessor represents the first argument".*

Em inglês, a expressão de dois argumentos por meio de uma expressão de possuidor é possível, já que, como se observou em (3-22b), a expressão dos argumentos pode ser feita por meio de genitivo e sintagma-de, respectivamente. Como visto em (3-22a), é vetada a expressão de argumentos por meio de expressões de possuidor do mesmo tipo (*of*, por exemplo). A gramática do português dificilmente licenciaria a expressão de duas ocorrências de sintagma-de para a expressão de dois argumentos em uma predicação nominalizada.

Como as nominalizações são construções encaixadas, elas podem atuar como termos em uma predicação de nível mais alto; em virtude dessa propriedade, são marcadas por uma interessante especificidade gramatical: não apenas as nominalizações podem manter praticamente intacta a estrutura argumental do predicado verbal *input*, preservando tanto a valência do predicado quanto as funções semânticas dos argumentos, mas podem atuar também como termos da predicação principal exercendo funções sintáticas, semânticas e pragmáticas que não seriam capazes de exercer caso fossem mantidos como verbos.

Do ponto de vista sintático, um falante pode selecionar um nome deverbal a fim de obter maior versatilidade de uso. Em português, por exemplo, uma oração subordinada finita é, sob a influência da LIPOC[8] (Dik, 1989), condicionada a adotar a posição final da predicação, ou, mais raramente, pode aparecer na posição P1. A predicação nominalizada, sob a forma de um SN, oferece menor complexidade categorial, podendo aparecer na posição de sujeito (3-23a), de objeto (3-23b), depois de preposições (3-23c) etc. (Mackenzie, 1996), o que não poderia ocorrer caso se mantivesse como um predicado verbal, por exemplo.

8 "LIPOC" é a sigla para "*language-independent preferred order of constituents*". Segundo esse princípio universal de ordenação, a preferência de colocação dos constituintes segue uma ordem de complexidade crescente, em que "complexidade" é definida nos seguintes termos: clítico < pronome < SN < sintagma adposicional < oração subordinada (cf. Dik, 1989, p.351).

(3-23) a *A colheita da soja deve ser apressada para aproveitar a alta do dólar.*
b *Pedro apressou a colheita da soja por causa da alta do dólar.*
c *Pedro dedicou-se à colheita de soja.*

Além disso, as nominalizações podem anteceder uma oração relativa, como em (3-24a), e permitir anáfora zero, como em (3-24b).

(3-24) a *A compra dos aparelhos que João providenciou não resolveu o problema.*
b *A construção de minha casa durou dois anos. Acredito que Ø não foi muito demorada, já que eu não tinha dinheiro para fazer mais depressa.*

Afirmar que as nominalizações induzem à redução da complexidade sintática é suficiente para entender a razão de selecioná-las. Um SN com um núcleo nominalizado é codificado por regras gramaticais mais simples e, por isso, dispõe de uma distribuição potencial geralmente não aberta à oração infinitiva correspondente.

3.2 A nominalização na Teoria da Gramática Funcional: preservação ou redução de valência

Discuti na seção anterior o tratamento da nominalização no âmbito da GF, discussão que se limitou ao que é mais geral e destituído de controvérsia. Esta seção será, todavia, devotada à discussão de aspectos mais particulares do assunto e também, certamente por isso, mais sujeitos à polêmica, como as diferentes posições a respeito de valência nas nominalizações.

Dik (1985; 1997) afirma que os termos de uma nominalização, definidos como primeiro e segundo argumentos, competem entre si pela posição de um sintagma de possuidor, competição que é, afinal, geralmente vencida pelo segundo argumento de uma predi-

cação transitiva. Reconhece, ainda, que a expressão de todos os argumentos sob a forma de possuidor é textualmente rara, visto que, dificilmente, nominalizações dispõem de mais de um argumento.

A posição de Dik (1985; 1997), indiretamente mantida em Keizer (2004a; 2004b), é que a representação subjacente da nominalização em geral preserva a valência quantitativa e qualitativa do verbo que lhe serve de *input*, conforme a regra de formação de predicados contida em (3-25) para nomes derivados por sufixação como *destruição*.

(3-25) Formação de nome deverbal em português (em –*ção*)
Input: pred [V] $(x_1)_{Agente}$ $(x_2)_{Meta}$
Output: pred-*ção* [N] $(x_1)_{Agente}$ $(x_2)_{Meta}$

(Adaptado de Dik, 1997, 165)

O esquema de predicado que serve de base para a nominalização está separadamente especificado no léxico, conforme representado em (3-26) e (3-27).

(3-26) Destruir $[_V]$ $(x_1)_{Agente/Força}$ $(x_2)_{Meta}$

(3-27) Destruição $[_N]$ $(x_1)_{Agente/Força}$ $(x_2)_{Meta}$

Mesmo tendo entrada lexical distinta do verbo, a nominalização mantém a correspondência sintático-semântica com o termo verbal por herança. Como na GF *standard* o esquema de predicado está fortemente associado com o item lexical correspondente, a diferença formal entre um predicado verbal e o predicado nominal derivado não está na estrutura argumental, que é a mesma, mas apenas no rótulo categorial que cada um recebe no léxico e nas diferenças mórficas de natureza flexional e derivacional.

Nessa perspectiva, é possível argumentar que, como a representação subjacente é de natureza abstrata, ela contém uma representação da estrutura argumental que pode diferir do uso na medida em que algum argumento não seja formalmente manifesto. Essa carac-

terização pode ser apenas inferida, já que Dik oferece apenas princípios formais de ajuste sem qualquer referência aos mecanismos funcionais, certamente presentes no uso, que poderiam bloquear o aparecimento de todos os argumentos. É sobre essa lacuna na formulação da Hipótese de Preservação de Valência que Mackenzie (1996, p.11) faz os seguintes comentários críticos:

> A desvantagem é que as regras de ajuste [...] são estipulativas e não fornecem nenhuma resposta à questão do motivo pelo qual as correspondências devem ser como aparecem. Além disso, não se dá nenhuma indicação das circunstâncias sob as quais um complemento se ajusta ou não: a funcionalidade específica das nominalizações [...] não se reflete no formalismo.[9]

O que seria possível acrescentar é que a redução das regras aos mecanismos formais de ajuste dá conta apenas da sintaxe, deixando de lado a prioridade das regras pragmáticas e semânticas, um dos postulados mais caros à GF. Mackenzie (1996) fornece como alternativa à proposta de Dik a Hipótese de Redução de Valência, para a qual predicados nominalizados são, assim como nomes lexicais plenos, avalentes. Reconhece, entretanto, que, na nominalização, a valência do predicado *input* perdura de algum modo. Para dar conta dessa propriedade, Mackenzie recorre, então, ao conceito de satélites implícitos, introduzido por Dik (1989).

A possibilidade de adicionar um satélite à predicação nuclear é determinada mais pelas propriedades combinatórias dos constituintes que a compõem do que pelo predicado em si. Dik argumenta que certos satélites de nível 1 já estão implícitos na predicação nuclear com base no fato de que predicações nucleares de natureza

9 Cf. o original: *"The drawback is that the adjustment rules [...] are stipulative, offering no answer to the question why the correspondences should be as they are presented. Furthermore, no indication is given of the circumstances under which a complement is adjusted or is not: the specific functionality of nominalizations [...] is not reflected in the formalism".*

controlada e/ou dinâmica implicam necessariamente que se desenvolvam de um determinado modo. Essa relação poderia ser descrita pela seguinte regra em (3-28):

(3-28) Dada uma predicação nuclear com os traços [+con] e/ou [+din], adicione uma posição estrutural $(x_i)_{Modo}$. Essa posição para Modo não será formalmente expressa a menos que ela seja suplementada com alguma informação mais específica (Dik, 1989, p.194).[10]

Mackenzie (1996) usa essa ideia em proveito da hipótese que defende em favor da redução argumental, o que envolve, por definição, o tratamento dos constituintes argumentais do verbo primitivo como satélites na nominalização, não sua preservação como argumentos.

Com base na noção de satélite implícito e o modo de notação proposto por Dik (1989) para esses dois tipos de termos (x para argumento e y para satélites), Mackenzie (1996, p.12) sugere uma alteração nessa distinção, explorando, pelo menos para entidades de primeira ordem, todas as letras finais do alfabeto: (x) para argumentos, (y) para satélites implícitos e (z) para satélites não implícitos. Satélites-y são tratados como quase argumentos, isto é, são considerados tão essenciais para as predicações dinâmicas e controladas quanto os argumentos. Satélites-z, por outro lado, não seriam automaticamente gerados, tendo, por isso, de ser adicionados como formas opcionais, do modo exatamente previsto pela teoria então vigente de satélites.

Considere-se a gradação entre construções completamente verbais e construções completamente nominais do inglês como em (3-29a-d).

10 Cf. o original: *"Given a nuclear predication with the features [+con] and/or [+din], add a slot* $(x_i)_{Manner}$*; this Manner slot will not be expressed, unless it is further in with some more specific information".*

(3-29) a *That* ***my horse*** (A^1) *won* ***the race*** (A^2) *came as a great surprise*. (oracão finita)

b ***My horse*** (A^1) *winning* ***the race*** (A^2) *came as a great surprise*. (gerúndio)

c ***My horse's*** (satélite-y) *winning* ***the race*** (A^2) *came as a great surprise*. (gerúndio-genitivo)

d ***My horse's*** (satélite-z) *winning* ***of the race*** (satélite-y) *came as a great surprise*. (nominalização produtiva)

Com base nela, Mackenzie (1996) argumenta que, enquanto (3-29a) é uma predicação verbal completa, (3-29b) mostra deverbalização parcial, retendo a valência; já (3-29c) consiste num caso de redução de valência: enquanto se preserva o segundo argumento (*the race*), o primeiro pode ser omitido ou expresso sob a forma do que denomina "satélite de possuidor" (*my horse's*), conforme a regra em (3-30):

(3-30) Formação do gerúndio-genitivo (do inglês)

Input: $\text{Pred}_V \ \text{Arg}_1 \ \text{Arg}_2$

Output: $\text{Pred}_{VN} \ \text{Arg}_2$

A nominalização produtiva contida em (3-29d) é segundo Mackenzie (1996) um caso de redução total de valência, já que os dois argumentos podem ser tanto omitidos quanto especificados por meio do satélite de Possuidor. Por isso, Mackenzie (1996) postula a regra contida em (3-31) no lugar da regra contida em (3-30).

(3-31) Nominalização produtiva (do inglês)

Input: $\text{Pred}_V \ \text{Arg}_1 \ \text{Arg}_2$

Output: Pred_n

Para Mackenzie (1996), a especificação dos argumentos do predicado *input* é feita por meio da expressão de satélites, no caso *my horse's* e *the race*, em (3-29d). A proposta de Mackenzie para o gerúndio-genitivo de (3-29c) (*my horse's winning the race)* é consi-

derar o Possuidor (*my horse's*) como um satélite-y, já que o papel de agente parece estar implícito: não se pode sensatamente falar sobre "*winning the race*" sem implicar que alguém ou alguma coisa venceu a corrida, ou a está vencendo ou ainda vai vencê-la.

Agora, no caso de nominalizações produtivas do inglês como (3-29d), que contêm dois termos sob a forma de possuidor, *my horse's* e *of the race*, Mackenzie (1996) propõe que o termo equivalente ao segundo argumento do verbo *input win* seja um satélite-y e que o equivalente ao primeiro argumento seja um satélite-z. Justifica essa classificação o fato de que equivalentes A^2 do predicado verbal *input* têm uma relação mais íntima com predicados nominalizados do que equivalentes de A^1 (cf. Mackenzie, 1996, p.14).

Satélites-z são adicionados somente quando há razão comunicativa para tanto, nos casos em que eles estão focalizados. Em nominalizações do inglês, equivalentes de A^2 podem aparecer ou como modificadores de posse pós-nucleares (sintagmas-de) ou como modificadores de posse pós-nucleares (genitivos). Se for focalizado um equivalente de A^2 na função de modificador de posse pré-nuclear, é gramaticalmente inaceitável adicionar um termo equivalente a A^1 na posição pós-nuclear como é ilustrado pelas diferenças entre (3-32) e (3-33).

(3-32) a *Caesar's (z) destruction of the city (y).*
b *CAESAR's destruction of the city.*

(3-33) a *The city's (y) destruction by Caesar (z).*
b **THE CITY'S destruction by Caesar.*

A proposta de Mackenzie prevê, então, que, se um termo equivalente a um argumento do verbo *input* corresponder a um satélite-y, ele pode ser omitido por estar implícito, mas é quase tão essencial à predicação quanto os argumentos. Agora, se um termo equivalente a um argumento corresponder a um satélite-z, ele é não implícito e, portanto, a inserção ou não do termo correspondente é apenas uma opção do falante. Uma razão teórica para representar

participantes opcionalmente expressos como satélite nas construções derivadas é, segundo Keizer (2004b, p.11), que:

> não há necessidade de uma regra separada de formação de termo-predicado, introduzida por Mackenzie e Hannay (1982 apud Keizer, 2004b) para tornar possível que termos indicando, por exemplo, locação ou posse funcionem como restritores sem serem antes convertidos em predicados. Para que esses termos funcionem como restritores eles precisariam primeiramente ser convertidos em predicados. Se, por outro lado, eles funcionarem como satélites, tal conversão não se tornaria mais necessária.[11]

Esse problema não mais se aplica, já que no arcabouço teórico da GDF usam-se variáveis separadas para representar categorias semânticas como locação e tempo (cf. Hengeveld; Mackenzie, 2006).

Outra vantagem teórica dessa proposta é que ela prevê uma correlação formal entre a estrutura argumental e a progressiva deverbalização do predicado *input:* a estrutura argumental de construções completamente verbais, como (3-29a), é progressivamente reduzida, até a ausência total de valência nas construções completamente nominais, como (3-29d). O processo de redução de valência serve de parâmetro para comprovar o processo gradual de descategorização dos verbos e a consequente recategorização deles como nomes.

A superioridade desse tratamento, segundo Mackenzie (1996, p.15), está no fato de que as regras de formação de predicado deixam de ser apenas estipulativas, conforme se vê na hipótese de preservação de valência de Dik, com a vantagem de a funcionalidade da nominalização, caracterizada por sua compactividade, estar

---

11 Cf. o original: *"there is no need for a separate rule of term predicate formation rule (Mackenzie and Hannay 1982 apud Keizer 2004b). This rule was introduced to make it possible for terms indicating e.g. location or possession to function as restrictors. For these terms to function as restrictors they would first need to be converted into predicates. If, on the other hand, they function as satellites, such conversion is no longer needed".*

diretamente refletida na noção de redução de valência: desprovê-las de seus "acessórios" é uma vantagem da qual dispõe o falante e justamente por isso é que são frequentemente empregadas.

Embora pareça tipologicamente instigante, a hipótese de Mackenzie apresenta alguns problemas que são derivados, a meu ver, da tendência de lidar com evidências baseadas somente em dados intuitivos sem qualquer consideração de outros dados reais de uso. Isso obscurece, por exemplo, sua visão da atribuição de saliência pragmática a certos constituintes do discurso, isto é, das razões pelas quais um termo necessita ser expresso por consistir em informação nova sem a qual o ouvinte não identifica o referente pretendido. Parece, com efeito, descabido tratar esses constituintes como satélites-y.

Além disso, no caso de um predicado nominal bivalente em que os dois termos representam informação comunicativamente necessária, parece não ser correto deixar de tratar A^1 como um satélite-z, que é uma informação adicional, para passar a interpretá-lo como satélite-y. Não haveria, assim, espaço para a possibilidade de caracterizar os termos de forma diferente mediante diferentes necessidades comunicativas.

Mackenzie considera que a dupla ocorrência da mesma função semântica em uma predicação não é problemática, dado o *status* diferencial dos dois possuidores como satélites implícitos e não implícitos. Segundo Dik (1989, p.103), esquemas de predicado nucleares nunca contêm mais que uma instância de uma dada função semântica. A afirmação de Dik está de acordo com o critério de atribuição de papel semântico proposto pela Gramática Gerativa: segundo o critério-theta (atribuição de papel temático), um constituinte só comporta um único papel temático e um mesmo papel temático só pode ser atribuído a um único constituinte (cf. Raposo, 1992). Considerando, por exemplo, a função semântica de Agente da GF, é impossível atribuí-la a mais de um argumento numa predicação verbal.

Para evitar a atribuição da mesma função semântica de Possuidor para os dois argumentos da nominalização, Mackenzie postula os conceitos de satélite implícito e de satélite não implícito, mas a

validade desses conceitos ficaria restrita ao âmbito específico de casos de nomes genéricos, porque em nomes não genéricos, como, por exemplo, o nome *destruction* em *Caesar's destruction of the city*, ambos os argumentos destacados aparecem explícitos.

Além disso, entende Mackenzie (1966) que os termos da nominalização são apenas metaforicamente possuidores, não podendo, assim, ser tratados como argumentos aos quais se atribuem funções semânticas de Agente e Meta (Paciente), por exemplo. Ao contrário, as funções semânticas de um predicado nominal de ação [+ dinâmico, + controlado] são Agente e Meta (Paciente), mas, para a conceitualização como objeto, é necessário transformá-las em algo formalmente relacionado a nomes e não a verbos.

Os nomes prototípicos contêm, de fato, possuidores, como *de Maria* em *o livro de Maria*. Nominalizar um predicado verbal é aproximá-lo da estrutura de um nome, ajustando os termos do predicado à estrutura prototípica do nome. Não se trata, assim, de uma verdadeira função semântica de Possuidor, mas de termos morfossintaticamente relacionados ao nome como se fossem sintagmas expressando função de Possuidor.

Num trabalho mais recente (Mackenzie, 2004a), a função semântica de Possuidor foi substituída por uma função Associativa, que daria conta genericamente das variadas funções semânticas que exercem os sintagmas preposicionais com *de* como modificador tanto de núcleos referindo-se a entidades de primeira quanto a de segunda ordem. Persiste assim mesmo o problema de tratar diferentes relações gramaticais num mesmo rótulo geral.

Apesar da generalização que ela contém, a hipótese de redução de valência não é destituída de problemas. É possível alistar pelo menos cinco pontos críticos:

1. Não está claro como tratar, por exemplo, a saliência pragmática de certos termos do discurso, isto é, como explicar que, em algumas construções, um termo representando informação nova deva ser expresso, mesmo se o destinatário não for capaz de identificar o referente dele. Esses termos deveriam

ser tratados como satélites-y? Além disso, num predicado nominal bivalente em que os dois argumentos contêm informação igualmente requerida por razões pragmáticas, o primeiro argumento deveria deixar de ser um satélite-z para tornar-se um satélite-y?[12]

2. A hipótese de redução de valência parece ser ainda mais problemática, se vista sob o prisma do processamento cognitivo, caso se assuma a correspondência intuitiva entre o predicado verbal *input* e o nominal derivado. Parece ser contraintuitivo reduzir a valência no processo de nominalização para, em seguida, adicionar satélites referindo-se aos argumentos potenciais. Em outros termos, por que reduzir os argumentos para voltar a recuperá-los como satélites?[13]
3. Mackenzie (1996) emprega a noção de satélite-z, como um atalho para qualquer satélite não implícito, para formular a hipótese de redução de valência. Entretanto, a categoria de satélite-z foi criada por Dik (1989) para contrastar com

12 Como a distinção entre satélite-y e satélite-z pode ser considerada puramente semântica em natureza, um satélite de qualquer um desses tipos pode ser comunicativamente saliente. De fato, na perspectiva de Mackenzie (1996), satélites estão tipicamente ausentes e são introduzidos somente se forem comunicativamente salientes. No entanto, como em meu enfoque a valência nominal é uma propriedade já prevista pelo item selecionado do léxico, é o estatuto pragmático de informação dada, ou Tópico Dado, ou da informação nova, ou focal, que se aplica à estrutura argumental, posição que rejeita, de pronto, as noções de satélites-y e satélite-z.

13 Como Mackenzie (1996) assume um esquema avalente para nomes derivados no Fundo, seu enfoque não incorpora a operação de redução de valência precedendo a anexação de satélites. A atribuição de contraintuitividade depende do modo como se vê o léxico; em minha abordagem, deve haver uma correspondência semântica entre o nome derivado e o verbo *input* e consequentemente um esquema correspondente entre as duas categorias lexicais. No entanto, numa abordagem como a de Mackenzie esse problema de processamento dependeria também do número de casos em que os argumentos são expressos comparados ao número de casos em que eles não são. É verdade que, se na maioria dos casos os argumentos não são expressos¸ o esforço cognitivo envolvido em adicioná-los como satélites seria muito menor que o esforço envolvido em processar que certas posições argumentais não são preenchíveis.

a de satélite-y, com o objetivo de cobrir a noção de satélite implícito, um satélite que está sempre previsto em estados de coisas dinâmicos e controlados. O problema é que a adição de uma nova categoria vai frontalmente de encontro ao princípio metodológico da economia, dado o fato de que seria possível fazer uma análise exaustiva dos fenômenos linguísticos envolvidos com as categorias já disponíveis de argumento e satélite. A mesma crítica pode ser estendida à hipótese de redução de valência.

4. Para justificar a hipótese de redução de valência, Mackenzie (1996, p.11) considera que predicados completamente nominalizados são, como nomes lexicais, avalentes. No entanto, é sobejamente conhecido que são monovalentes muitos nomes lexicais não derivados como *medo do escuro*, *vontade de doce*, *ciúme da namorada* etc. Nada mais teoricamente justo, portanto, que considerar os nomes deverbais como valenciais em virtude da evidente relação que mantêm com os predicados verbais correspondentes como, por exemplo, *destruir a cidade* > *destruição da cidade*.
5. A aplicação dos rótulos satélite-y e satélite-z varia conforme a relevância topical ou focal do termo na estrutura interna do nome deverbal.

Acredito que essas questões podem receber uma resposta apropriada com base na análise de exemplos genuínos de nominalizações extraídos de uma amostra de dados de conversação real submetida a um procedimento estatístico.

Uma tarefa importante, portanto, é proceder ao processamento e à análise dos dados e, em seguida, reformular a hipótese de preservação de valência de um modo tal que ela se torne não apenas mais funcionalmente eficiente, mas também mais alinhada aos princípios gerais da GDF.

Antes de proceder a esse trabalho empírico, reservado para a Parte II, pretendo desenvolver, na seção seguinte, a discussão do modo como a nominalização é tratada na GDF. Como, todavia, essa

versão refaz alguns conceitos da GF, é necessário explicitar, inicialmente, alguns aspectos da nominalização na GF para, em seguida, ver como foram finalmente repropostos pela GDF.

3.3 A nominalização na Gramática Discursivo-Funcional

Considerem-se as construções contidas em (3-34a) e (3-34b).

(3-34) a *A Prefeitura construiu a ponte.*
b *A construção da ponte pela Prefeitura.*

Trata-se de duas construções formalmente diferentes, já que (3-34a) é nucleada em torno de um verbo, enquanto (3-34b) é nucleada em torno de um nome. Entretanto, essas diferenças não são capazes de obscurecer a interpretação de que as duas construções representam basicamente o mesmo estado de coisas; além disso, seus predicados têm restrições de seleção idênticas e podem ser usados em circunstâncias similares. Observadas essas características similares, o problema básico que se coloca para a teoria da linguagem é como fornecer uma descrição estrutural que seja capaz de preservá-las formalmente.

No âmbito da Teoria Gerativa, a solução transformacional na chamada Teoria *Standard* era uma derivação (Raposo, 1992). Supunha-se, na ocasião em que predominava esse modelo, que as nominalizações como a de (3-34b) seriam deriváveis da representação subjacente de sentenças completas como (3-34a) por meio de uma regra transformacional de nominalização. Esse tipo de descrição enfatizava as similaridades das duas construções em prejuízo de suas diferenças.

Essa solução transformacional, que era à época postulada pela versão padrão da Gramática Gerativa, nunca foi admitida pela GF, que, por princípio, evita transformações, com base na suposição de que, uma vez construída, a estrutura é mantida em toda represen-

tação adicional que se faz da expressão linguística. A derivação é, assim, uma questão de expansão gradual em vez de uma projeção transformacional de uma estrutura sobre outra.

Há, todavia, um componente do modelo da GF que, de fato, requer uma regra de derivação, o de formação de predicados no léxico. Regras de formação servem para derivar um esquema de predicado de outro, estendendo, assim, o conjunto de esquemas de predicado disponível no léxico para um conjunto de esquemas de predicado derivados, que podem, então, fornecer o *input* necessário para a construção de predicações.

Alguns tipos de regras de formação de predicados podem realizar mudanças no estatuto categorial do esquema *input*, reduzindo ou estendendo o número de argumentos, incorporando argumentos ao predicado, ou modificando as funções semânticas das posições de argumento do esquema de predicado.

Essas regras de derivação são, contudo, confinadas ao "fundo" da gramática no arcabouço da GF, o componente que contém o conjunto de predicados e o conjunto de termos a partir dos quais predicações podem ser construídas (Dik, 1989, p.19-21). Se essas regras de derivação podem ser consideradas um tipo de transformação, elas têm muito pouco a ver com as regras transformacionais típicas do modelo *standard* da Gramática Gerativa, que eram de natureza sintática e não lexical como são as regras de formação de predicado na GF.

A partir da segunda metade da década de 1960 e especialmente durante os anos 1970, os gerativistas procuram reduzir a capacidade descritiva da teoria com o subsequente aumento da capacidade explicativa e redução de seu poder expressivo mediante a simplificação do formato das regras. Uma das estratégias mais notórias consistiu em extrair da formulação de regras específicas e particulares os aspectos que podiam ser convertidos em princípios gerais da linguagem (cf. Raposo, 1992, p.52). O trabalho teórico operado nesse período aboliu a quantidade excessiva de regras, reduzindo-as a um pequeno conjunto de opções que atuam como princípios restritivos universais na caracterização de gramáticas possíveis.

A redução de regras categoriais de reescrita, operada por Chomsky (1970) mediante os princípios da Teoria X-barra (cf. Raposo, 1992), incluiu a abordagem derivacional da nominalização, que foi abandonada em favor de uma descrição lexical. Na redução da capacidade descritiva das regras transformacionais, a nominalização deixou de ser tratada transformacionalmente e passou a ser caracterizada através de uma interação do léxico com o sistema de regras de reescrita. Passou a predominar a ideia de que o verbo *construir* e o nome *construção* entram no léxico como representantes verbal e nominal de um item lexical comum. Uma implicação dessa posição é que a nominalização não é mais derivada da representação subjacente de uma sentença completa.

Por essa razão, surgiu a necessidade de um mecanismo para formular regularidades por meio de categorias sintáticas de modo a capturar as propriedades comuns de construções verbais e de construções nominais. A solução para esse problema, a já mencionada Teoria X-barra, permitiu que as regularidades sintáticas pudessem ser formuladas em termos da categoria variável X, que pode assumir valores como verbo (V), nome (S), adjetivo (A) ou preposição (P).

Problemas similares também surgiram no quadro da GF; dadas, porém, as posições teóricas básicas desse modelo, existem certas diferenças relevantes, que cabe enumerar.

Em primeiro lugar, um princípio da GF diz que todos os predicados dotados de conteúdo lexical que não podem ser derivados de outros predicados por regra produtiva devem ser alocados no léxico. Não existem regras produtivas no português como a do gerúndio em inglês, que deriva um nome de qualquer verbo, como *constructing* do verbo *construct*. Em português, por exemplo, diferentes verbos dispõem de diferentes processos de derivação, como se comprova com a formação de *construção* a partir do verbo *construir*, e *constrangimento* a partir do verbo *constranger*. Deduz-se disso que nomes desse tipo devem ter sua própria entrada no léxico. Assim, o critério geral de produtividade faz cumprir a solução lexical para o problema.

Em segundo lugar, o fato de terem os itens *construir* e *construção* entradas distintas no léxico não implica a necessidade de uma nova

teoria para considerar as similaridades entre eles. Cada predicado é parte de um esquema de predicado, e não há nada que previna o fato de predicados nominais terem esquemas de predicado muito similares ou mesmo idênticos ao esquema de predicado de verbos relacionados. Nesse caso particular, postulam-se os esquemas de predicado para *construir* e *construção* contidos, respectivamente, em (3-35) e (3-36).

(3-35) construir [V] $(x_1)_{\text{Agente/Força}}$ $(x_2)_{\text{Meta}}$

(3-36) construção [N] $(x_1)_{\text{Agente/Força}}$ $(x_2)_{\text{Meta}}$

Também o fato de *construir* e *construção* terem argumentos com as mesmas funções semânticas está plenamente representado nesses esquemas de predicado. Além disso, a relação entre esses predicados pode ser expressa na definição de significado de *construção* contida em (3-37).

(3-37) construção [N] $(x_1)_{\text{Agente/Força}}$ $(x_2)_{\text{Meta}}$
$=_{\text{df}}$
ação/evento definido por
construir [V] $(x_1)_{\text{Agente/Força}}$ $(x_2)_{\text{Meta}}$

O esquema de predicado para *construção* pode ser usado em formação de termos de acordo com o modelo de uma construção encaixada, como se observa em (3-38).

(3-38) (d1e_i: [construção [N] $(\text{prefeitura})_{\text{Agente}}$ $(\text{ponte})_{\text{Meta}}$])[14]

Ao determinar a expressão formal dos argumentos dentro do contexto de uma nominalização, as regras de expressão produzem

14 Nesse esquema de predicado, **d** representa "definido", **1** representa "singular", e **e,** "estado de coisas".

alternativas como as seguintes ocorrências: *a construção da ponte pela Prefeitura* e *a construção pela Prefeitura da ponte*.

Uma propriedade distintiva da GF é que, ao lado da derivação sintática completa e da listagem lexical pura, há também o mecanismo das regras produtivas de formação de predicado. Assim, em uma língua como o inglês, em que nomes podem ser produtivamente derivados de verbos correspondentes, as relações relevantes podem ser expressas em uma regra de formação de predicado.

Garcia Velasco e Hengeveld (2002) discutem o papel dos esquemas de predicado dentro da GF e apresentam a possibilidade de substituí-los por esquemas de predicação mais gerais, no interior dos quais os lexemas, listados no léxico sem esquemas, são inseridos por meio de aplicação de regras de ligação. A ideia de introduzir esquemas de predicação na GF foi proposta originalmente por Hengeveld (1992 apud Garcia Velasco; Hengeveld, 2002) como uma alternativa para a Regra de Formação Termo-Predicado de Dik (1980 apud Garcia Velasco; Hengeveld, 2002).

Os autores argumentam que a noção de esquema de predicado deve ser substituída, por um lado, por uma combinação de esquemas de predicação e, por outro, por lexemas fornecidos com definições de significado abstrato; finalmente, deve haver um mecanismo de ligação que permita conectar esses dois elementos na gramática. Os autores defendem o princípio de que essa passagem não é somente desejável, mas também ajuda a teoria a atingir um grau mais elevado de adequação psicológica, tipológica e pragmática.

O modelo de adequação psicológica requer que a GF possa referir-se tão estreitamente quanto possível aos modelos de competência linguística e comportamento linguístico (Dik, 1997, p.13). Esquemas de predicado desempenham um papel central na organização da GF, dado que eles constituem o *input* para a geração oracional e contêm uma grande quantidade de informação sintática e semântica relevante ao processo interpretativo. Espera-se, então, que a evidência psicolinguística confirme seu estatuto privilegiado na produção e na compreensão de sentença. Um caso interessante para avaliar a adequação psicológica dos esquemas de predicado

concerne ao comportamento de verbos que permitem estruturas de argumento alternativas, como, por exemplo, os que participam das chamadas alternâncias causativas, como ilustradas em (3-39):

(3-39) a *Pedro abriu a porta.*
b *A porta abriu.*

(Adaptado de Garcia Velasco; Hengeveld, 2002, p.97)

Conforme discutido na seção anterior, a GF trata essas alternâncias mediante a formulação de Regras de Formação de Predicado. De uma perspectiva psicolinguística, esse tratamento sugere que a interpretação e a produção de um esquema *output*, como o contido em (3-40), terão que se referir à regra de formação de predicado relevante para estabelecer a relação com o esquema *input* e chegar à interpretação semântica apropriada do predicado derivado:

(3-40) REGRA DE FORMAÇÃO INCOATIVA
INPUT: $\text{pred}_V\ (x_1)_{\text{Agente}}\ (x_2)_{\text{Meta}}$
OUPUT: $\text{pred}_V\ (x_2)_{\text{Processado}}$

(Adaptado de Garcia Velasco; Hengeveld, 2002, p.97)

Verbos que atuam em alternância causativa, como os de (3-39a-b), não apresentam diferentes sentidos, mas diferentes estruturas de argumento. Ao se deparar com um verbo desse tipo, o falante terá que "tomar uma decisão" para determinar a que esquema de predicado se vincula o verbo. A previsão, que se origina da organização do léxico na GF, é, primeiramente, que a sentença (3-39b) poderia ser mais complexa para processar que (3-39a), uma vez que ela requer a abertura da regra de formação de predicado relevante para chegar à interpretação correta do predicado.

Para iniciar a discussão da adequação pragmática, os autores mencionam o exame que Butler (1998 apud Garcia Velasco; Hengeveld, 2002) faz do conceito de esquema de predicado. Para Butler, o modelo de relação predicado-argumento, resumido no esquema

de predicado, não é capaz de levar em conta o uso do predicado e de seus argumentos na situação discursiva (Butler, 2001, p.56 apud Garcia Velasco; Hengeveld, 2002). Examinando o uso do predicado verbal *give* (dar), Butler conclui que a estrutura apresentada em (2-7), repetida abaixo como (3-41) por conveniência, é simples demais para levar em conta o uso real do verbo.

(3-41) Pass [[dar (João) (o livro) (a Maria)] (na biblioteca)]]

(Dik, 1989, p.68)

Com efeito, em quase 20% dos exemplos, o agente não é animado como o esquema requer; o exemplo fornecido pelo autor está contido em (3-42).

(3-42) *Anos compondo versos gregos deram-lhe confiança.*

(Adaptado de Butler, 2001, p.58 apud Garcia Velasco; Hengeveld, 2002, p.101)

A alta incidência de agentes não animados sugere, para os autores, que esse exemplo de Butler constitui não apenas uma violação às restrições de seleção do primeiro argumento, mas tem também implicações mais sérias.

A GF define a função semântica de Agente como a entidade controladora de uma Ação, e, por isso, todos os testes usados por Dik para identificar o tipo semântico de um estado de coisas[15] mostram que o primeiro argumento do exemplo contido em (3-42) deve ser claramente identificado como não agente. Isso acarreta a possibilidade de atribuir a esses argumentos inanimados uma função semântica não agentiva, e o candidato mais provável é a função Força, que representa uma entidade não controladora, mas que é responsável pelo desencadeamento do estado de coisas.

---

15 Os testes compreendem a predicação ser complemento de verbos como *ordenar*, *persuadir* e *requerer* e a predicação combinar com satélites na função de Beneficiário.

Essa alternância entre Forças e Agentes na posição de primeiro argumento não se restringe ao verbo *dar*. Tem sido reconhecido, já de longa data, que satélites com função de Instrumento podem assumir a posição de sujeito, um processo que, na GF, acarreta a conversão de Instrumento em Força, como em (3-43a-b). Dada a produtividade dessa alternância, os autores consideram a viabilidade de formular uma Regra de Formação de Predicado dentro do arcabouço da GF com o formato geral contido em (3-43c):

(3-43) a *João cortou a carne com uma faca.*
b *A faca cortou a carne.*
c REGRA DE FORMAÇÃO INSTRUMENTO-FORÇA
INPUT: $(f_1\text{: pred}_V\ (f_1)\text{: }\{(f_2\text{: }(x_3)_{\text{Instrumento}}\ (f_2))\}\ (f_1))\ (x_1)_{\text{Agente}}\ (x_2)_{\text{Meta}}$
OUTPUT: $(f_1\text{: pred}_V\ (f_1))\ (x_3)_{\text{Força}}\ (x_2)_{\text{Meta}}$
SIGNIFICADO: Entidade instrumental (x_3) provoca o processo denotado por pred_V

(Adaptado de Garcia Velasco; Hengeveld, 2002, p.102)

No entanto, segundo o entendimento dos autores, em termos de processamento cognitivo é excessivamente elevado o custo teórico de regras como a contida em (3-43c) pelas seguintes razões: em primeiro lugar, ela causa uma alteração radical no esquema de predicado derivado, que, de qualquer maneira, oculta a relação semântica óbvia entre as duas variantes; em segundo lugar, ela nos obriga a aceitar que regras de formação de predicado podem admitir satélites no *input*, o que constituiria um postulado um tanto controverso em virtude de a aplicação de regras de formação de predicado ser restrita ao Fundo, que é lexical por natureza.

Quanto ao princípio de adequação tipológica, Garcia Velasco e Hengeveld (2002) mencionam a tipologia de sistemas de classes de palavras ou partes do discurso de Hengeveld (1992), que usa definições funcionais, isto é, definições distribucionais para as classes de palavras. De acordo com essas definições, é possível distinguir

classes de palavras com base nas posições sintáticas que ocupam, como núcleo ou modificador de um sintagma de termo ou de um sintagma de predicado.

O resultado da aplicação dessas definições a uma grande diversidade de línguas mostra que algumas delas dispõem de uma classe de lexemas somente em uma dessas posições sintáticas, enquanto outras permitem o uso de uma única classe de lexemas em várias posições sintáticas. Na terminologia de Hengeveld (1992), as línguas do primeiro tipo consistem num sistema *rígido* de classes de palavras, enquanto as do segundo tipo consistem num sistema *flexível*. Línguas com sistemas mais flexíveis de classes de palavras usam uma classe única de lexemas para as diferentes posições sintáticas.

Em línguas tipologicamente flexíveis, os esquemas de predicado não são uma propriedade intrínseca dos lexemas em si; ao contrário, os lexemas estão associados a um esquema de predicado que licencia a ocorrência deles conforme a necessidade da situação de comunicação. Segundo os autores, para fazer justiça a esse aspecto de línguas flexíveis, é necessário separar os contextos sintáticos dos lexemas. Se línguas flexíveis requerem uma abordagem como essa, o mesmo se aplica a todas as outras línguas como uma decorrência da necessidade de adequação tipológica da GF (Garcia Velasco; Hengeveld, 2002, p.107).

Um modo de resolver o problema que a noção de esquema de predicado apresenta com respeito aos padrões de adequação psicológica, pragmática e tipológica é dividir esquemas de predicado em predicados e esquemas, e ter um mecanismo de ligação que os associe na gramática. Assim, os autores dão preferência por substituir a expressão *esquema de predicado,* empregada por Dik (1989), por *esquema de predicação*, e o termo *predicado*, também empregado por Dik, por *lexema.* Desse modo, o termo *predicado* torna-se disponível para qualquer tipo de item lexical, ou "lexema", usado como predicados em contextos sintáticos particulares; fica reservada para a expressão *esquemas de predicação* a função de especificar as configurações subjacentes em que podem ocorrer os lexemas. Nesse

caso, o conjunto de esquemas de predicação não é idêntico para todas as línguas, nem sintática nem semanticamente.

Ao definir os esquemas de predicação, Garcia Velasco e Hengeveld (2002) empregam o formalismo desenvolvido por Hengeveld (2004d),[16] mediante o qual é possível estabelecer uma distinção básica entre unidades no Nível Interpessoal e Representacional, respeitando a nova versão da GF, a GDF.

Conforme discutido na Seção 2.5, as unidades do Nível Interpessoal são definidas em termos de sua função na comunicação. As variáveis usadas para formalizá-las são letras maiúsculas. As funções relevantes aqui são Subatos de Referência (*reference*) (R) e de Atribuição (*ascription*) (T). Em muitas outras abordagens, esta última função seria chamada de predicação, mas esse termo é reservado, na GDF, para uma instanciação semântica, no Nível Representacional, de um Subato de Atribuição, definido, portanto, no Nível Interpessoal.

Unidades no Nível Representacional são definidas em termos do tipo de entidade que designam. As variáveis usadas para representá-las formalmente são letras minúsculas. Os tipos de entidade relevantes desse nível são propriedades/relações (f), indivíduos (x), lugar (l), tempo (t), estados de coisas (e), episódio (ep) e conteúdo proposicional (p). Assim, mediante o uso de unidades de diferentes níveis, separam-se rigorosamente funções comunicativas e categorias semânticas.

Os dois níveis de análise são em princípio independentes um do outro. Todos os tipos de entidade podem ser referidos, e todos os tipos de entidade podem ser predicados. Assim, um Subato de Referência (R) pode fazer uso da designação de um indivíduo (x), como em (3-44a), bem como da descrição de um estado de coisas (e), como em (3-44b).

---

16 À época da publicação de Garcia Velasco e Hengeveld (2002), Hengeveld (2004) é mencionada como obra no prelo, o que explica a posterioridade cronológica da fonte citada.

(3-44) a *Ele viu o menino.* $(R_1: (x_1) (R_1)$

b *Ele viu o menino fazer um desenho.* $(R_1: (e_1) (R_1)$

(Adaptado de Garcia Velasco; Hengeveld, 2002, p.110)

Da mesma forma, o inverso também é verdadeiro: pode-se usar a descrição de uma entidade de primeira ordem num Subato de Referência, como em (3-45a), e igualmente também num Subato de Atribuição, como em (3-45b).

(3-45) a *Estou procurando um carpinteiro.* $(R_1: x_1 (R_1)$

b *Meu vizinho é carpinteiro.* $(T_1: x_1) (T_1)$

(Adaptado de Garcia Velasco; Hengeveld, 2002, p.110)

As variáveis assim distinguidas permitem formular esquemas de predicação como o de (3-46a) para um tipo intransitivo de predicação e o de (3-46b) para um tipo transitivo sem levar em conta as categorias semânticas envolvidas com o predicado, que já vêm definidas do léxico. Para dar conta da inserção lexical, o símbolo (◆) representa, no esquema de predicação, a posição estrutural para a realização dessa operação de ligação entre Léxico e Gramática.

(3-46) a Estado de coisas intransitivo: $(T_1: (f_1: \blacklozenge (f_1)) (T_1)) (R_1)$

b Estado de coisas transitivo: $(T_1: (f_1: \blacklozenge (f_1)) (T_1)) (R_1) (R_2)$

(Adaptado de Garcia Velasco; Hengeveld, 2002, p.110)

As formulações contidas em (3-47a-b) representam esquemas de predicação para a inserção de termo na função de núcleo de sintagma.

(3-47) a indivíduo: $(R_1: (x_1 \blacklozenge (x_1) (R_1) \ldots (R_N))$

b estado de coisas: $(R_1: (e_1 \blacklozenge (e_1)) (R_1) \ldots (R_N))$

(Adaptado de Garcia Velasco; Hengeveld, 2002, p.111)

O esquema de predicação de (3-47a) representa, então, um termo, como *paletó* em *o paletó novo dele*, enquanto o esquema de (3-47b), um estado de coisas, como *casamento* em *o terceiro casamento dele*. Assim, o esquema de predicação para uma nominalização representa a ideia de que um Subato de Referência, definido no Nível Interpessoal, tem a função de designar, no Nível Representacional, um estado de coisas, o que dá suporte ao caráter semanticamente complexo de nomes derivados. O formalismo, longe de ser estéril e vazio, é absolutamente relevante para uma representação adequada das nominalizações, como entidades que representam estados de coisas.

Como lexemas não se confundem com esquemas, um problema a resolver é como lexemas se ligam aos esquemas de predicação em que ocorrem. Na GF são os próprios esquemas de predicado que definem a valência quantitativa e qualitativa dos lexemas.

Ao contrário, quando se separam lexemas e esquemas de predicação, torna-se crucial desenvolver uma solução alternativa para o problema da ligação entre léxico e sintaxe, e a solução que Garcia Velasco e Hengeveld (2002) fornecem para a questão envolve o uso de primitivos, postulados para a definição de categorias ontológicas básicas de natureza conceitual, no mesmo espírito do que propõe Jackendoff (1990). São os componentes de significado abstrato que servem como gatilhos para a ligação entre o lexema envolvido e um esquema de predicação específico ou, dito em outras palavras, entre o Léxico e a Sintaxe. Se o mecanismo de ligação se baseasse em definições concretas de significado, cada elo teria que ser especificado separadamente.

Considerando os mecanismos abstratos de definição de significado, é possível considerar a definição, em (3-48b-c), para o lexema verbal *open*.

(3-48) a Open[V]
b [f_1: [CAUSE (x_1)[BECOME **open’** (x_2)]]]
c [f_1: [BECOME **open’** (x_1)]]

(Adaptado de Garcia Velasco; Hengeveld, 2002, p.116)

As definições de (3-48b-c) expressam que o lexema *open* designa uma relação representada pela variável **f** entre duas entidades, representadas pela variável **x**. Ao contrastar as duas definições de significado para o lexema *open*, fica evidente que a segunda definição consiste numa subparte da primeira. Essa relação significa que é tarefa teoricamente simples formular uma regra que derive um significado abstrato de outro lexema e, em seguida, ligar cada definição a um esquema de predicação.

A regra de formação de predicado passa a ser substituída pela regra de derivação de lexema, separando-se, assim, com maior nitidez, a sintaxe do léxico. Consequentemente, se um lexema tem mais de uma definição de significado, como o verbo *open*, ele pode ser vinculado a mais de um esquema de predicação. Essa solução será aplicada às nominalizações do português na Seção 5.3.

Uma forte motivação para essa separação entre lexemas e esquemas de predicado é o fenômeno da especificação ou não especificação de argumentos na função de objeto. Garcia Velasco e Portero Muñoz (2002) alegam que a não especificação de objetos em verbos como *comer* (*eu estava comendo*) é determinada por restrições de natureza lexical e de natureza discursiva. A diferença básica entre os dois processos é que a omissão lexical requer a ausência de um referente possível para o objeto subentendido. Fillmore (1986 apud Garcia Velasco; Hengeveld, 2002, p.103) observa que, no contraste entre *Ele estava comendo/Eu quero saber o que ele estava comendo* e entre *Eles descobriram/Eu quero saber o que eles descobriram*, não faz o menor sentido fazer a segunda pergunta, dado que o referente de *descobrir* deve ser suficientemente óbvio na situação discursiva.

Entendem Garcia Velasco e Portero Muñoz (2002) que é possível formalizar essa diferença, assumindo que os esquemas de predicado de Dik (1989), agora denominados esquemas de predicação, possam ser atribuídos *online* e os verbos que admitem omissão lexical assumem um esquema de predicado de um lugar em seu uso intransitivo, não permitindo, assim, um processo anafórico de recuperação do referente. Os autores postulam que esse processo pode ser pragmaticamente motivado, uma vez que ele licencia uma

alteração no tipo semântico de estado de coisas de um modo eficiente e econômico.

Este capítulo encerra a Parte I deste tabalho, que define aspectos diversos do posicionamento teórico em relação às classes de palavras. Na Parte II, procedo à análise de uma amostra de nominalizações do português falado para fornecer evidência para a hipótese de preservação de valência. O enfoque escolhido é, como se viu nas seções 3.2 e 3.3, o da GDF e, mais especificamente, no âmbito da GF, a hipótese de preservação de valência postulada por Dik (1985; 1989).

A análise empírica se limitará à demonstração de que a não especificação dos argumentos da nominalização não implica ausência de valência, como defende Mackenzie (1996), e à explicitação das motivações funcionais de natureza semântica e pragmática para a não especificação da estrutura argumental. Os aspectos teóricos e empíricos serão retomados na Parte III, que trata das generalizações e implicações teóricas em que essas motivações são interpretadas, agora, à luz da GDF.

Vale lembrar que o objetivo final é demonstrar que os nomes que representam entidades de ordem superior e, portanto, são providos de valência constituem membros não prototípicos da classe; ao contrário, os nomes não derivados e que representam entidades de primeira ordem e, portanto, são, por definição, desprovidos de valência, constituem membros prototípicos.

Parte II

A preservação da estrutura argumental na nominalização: evidências empíricas

4
A ESTRUTURA ARGUMENTAL NA NOMINALIZAÇÃO

Introdução

Este capítulo, que se destina ao exame da organização valencial na amostra selecionada, é o ponto de partida para a comprovação empírica da valência como propriedade constitutiva da nominalização, que constitui a tarefa principal da Parte II deste trabalho. O ponto-chave é a sustentação da hipótese de que há dois tipos de nominalizações: um deles se refere a entidades de primeira ordem e o outro se refere a entidades de ordem superior. A principal diferença entre eles é ausência e presença de valência respectivamente, que dá suporte à hipótese de continuidade categorial nas classes de palavras, tese que será desenvolvida na Seção 6.1, que trata dessa natureza complexa da nominalização.

O ponto de referência será a análise da estrutura argumental das nominalizações que representam entidades de segunda ordem em dados do português falado, porque comprovar que os nomes deverbais dessa natureza não são desprovidos de valência, ao contrário dos nomes de primeira ordem, é um ponto crucial deste trabalho.

A Seção 4.1 descreve alguns procedimentos metodológicos que presidem à seleção operada no universo de investigação e, ao mesmo tempo, faz a apresentação da amostra utilizada.

Antes de tratar especificamente da relação entre valência nominal e expressão argumental, vale a pena repetir que a ideia do *continuum* categorial que assumo neste trabalho está fortemente atrelada à sustentação da hipótese de preservação de valência das nominalizações, defendida por Dik (1985; 1997), conforme já mencionado na Seção 3.3. Segundo essa posição, nominalizações e verbos primitivos respectivos dispõem da mesma estrutura argumental, em virtude de haver uma correspondência estrita entre eles, baseada também na noção de herança argumental de Hoekstra (1986) entre predicados nominais derivados e os predicados verbais subjacentes, ainda que quase nunca os predicados nominais expressem abertamente a estrutura argumental de que dispõem por herança. Nesse caso, pretendo demonstrar, com dados reais de conversação em língua portuguesa, o modo como a valência nominal é preservada.

O critério de exposição escolhido é o seguinte: forneço, inicialmente, evidência supostamente favorável à hipótese de redução de valência apenas com base na interpretação de resultados mais gerais, para, em seguida, falseá-la com evidência francamente favorável à hipótese de preservação de valência mediante um tratamento mais detalhado e específico dos mesmos resultados em relação ao modo como o discurso se organiza nos Níveis Interpessoal e Representacional.

O percurso da análise é o seguinte: apresento, inicialmente, em 4.2 evidência em favor da ideia de competição argumental, mostrando que, no caso de predicados de dois lugares, a preferência é pela expressão por sintagma adposicional, introduzido pela preposição *de*, e o candidato derrotado na competição recebe expressão por anáfora zero.

Os poucos casos em que a expressão argumental se manifesta plenamente em predicados bivalentes são de nomes derivados de verbos com complemento oblíquo, que já dispõem de uma preposição, fenômeno que só reforça o princípio de competição. A especificação dos resultados torna-se ainda mais detalhada na Seção 4.3, em que trato da codificação morfossintática dos argumentos.

Na Seção 4.4, apresento evidência de que há casos de anáfora zero motivados por relações de predeterminação semântica entre a predicação encaixada e a predicação matriz, e casos de anáfora zero motivados por determinação pragmática em função do grau de acessibilidade dos referentes na memória de curto prazo dos participantes da interação. Demonstro também que há motivações semântico-pragmáticas para interpretar nominalizações em si mesmas como não referenciais, propriedade que determina, mesmo nos verbos, a ausência de expressão de valência, e para interpretar como indeterminada a referência de argumentos.

4.1 Apresentação da amostra e dos procedimentos de seleção

A amostra é constituída por 181 ocorrências extraídas de um *corpus* maior coletado por Santana (2005), incluindo 305 casos de nomes deverbais e deadjetivais, a partir de três inquéritos do Projeto NURC-SP, correspondentes a elocuções formais (EF-377: Castilho; Preti, 1986), diálogos entre informante e documentador (DID-237: Preti; Urbano, 1988) e diálogo entre dois informantes (D2-360: Castilho; Preti, 1987).

O primeiro consiste numa aula gravada, sendo, portanto, reduzido o grau de intercâmbio entre os polos da emissão e da recepção; o segundo e o terceiro são exemplos típicos de conversação, mas nos diálogos do tipo DID o intercâmbio é em geral controlado pelo documentador. Já o tipo D2, o mais informal dos três, consiste numa conversa livre entre dois informantes e, embora o documentador também esteja presente, ele tem participação extremamente reduzida na conversação. As pessoas entrevistadas têm escolaridade de nível superior; pode-se supor, por isso, que os três tipos de inquéritos são casos ilustrativos de fala padrão e relativamente formal.

O critério principal para levantamento da amostra consistiu na seleção de nominalizações mono e bivalentes de segunda ordem. O que esses predicados têm em comum com os predicados verbais

correspondentes é que eles descrevem eventos e estados que se caracterizam por serem localizáveis no tempo e serem avaliáveis em termos de seu estatuto real, traços que constituem critérios para a definição de estados de coisas (cf. Hengeveld, 2004c).[1]

Um critério para o levantamento desse tipo de nominalização de segunda ordem é a tipologia semântica de estados de coisas postulada por Dik (1989) com base nos traços [+/- dinâmico], [+/- controlado] e [+/- télico]. Como a tipologia tem apenas o efeito de comprovar a natureza superior da entidade representada pela nominalização, para reiterar seu caráter de estado de coisas, bastam os dois primeiros traços, procedimento que conduziu a uma tipologia da seguinte natureza: Posição [-dinâmico, +controlado], Estado [-dinâmico, -controlado], ação [+dinâmico, +controlado] e Processo [+dinâmico, -controlado].

Outro critério relevante é a natureza argumental do predicado nominal. Foram selecionadas as nominalizações derivadas de verbos monovalentes, bivalentes com complemento direto e bivalentes com complemento oblíquo, classificados na tradição gramatical como intransitivos, transitivos diretos e transitivos indiretos respectivamente.

Na realidade, a análise dos bivalentes oblíquos serve como parâmetro ou ponto de referência para a dos bivalentes com complemento direto para a verificação da hipótese de competição entre os argumentos pela posição de sintagma-de, já que, potencialmente, somente os nomes derivados de transitivos diretos seriam capazes de acionar a competição entre os argumentos para o preenchimento de uma posição única de possuidor. O argumento nominal, correspondente ao complemento oblíquo de predicados verbais bivalen-

1 É digno de nota, todavia, que as nominalizações aparecem não como construções independentes, mas como construções dependentes. Nesse caso, a análise das condições temporais e do estatuto de verdade depende dessas relações. Assim, por exemplo, em *João lamentou a demissão de Maria*, está claro que a predicação nominal é temporalmente anterior à predicação verbal em que se encaixa como complemento. Já o estatuto epistêmico de uma construção dependente é o mesmo de pressuposições em geral (ver a esse propósito Cristofaro, 2003).

tes, já dispõe de preposição, que é automaticamente herdada pela estrutura argumental da nominalização, fenômeno que, de saída, exclui a possibilidade de competição entre os argumentos sujeito e complemento.

É ocioso dizer que nenhuma amostra é capaz de representar todas as possibilidades de ocorrência da língua, tipo de problema metodológico que afeta esta pesquisa. Assim, numa primeira fase, o levantamento incluía os casos de predicados trivalentes; como não havia nenhuma ocorrência de nome derivado de verbo bitransitivo, e os poucos casos de trivalentes são predicados com A^3 na forma de oblíquo, fiz um recorte nessa segunda fase excluindo predicados desse tipo de estrutura argumental. Como tem relevância observar a competição entre os argumentos para ocupar a posição de sintagma de possuidor na nominalização, incluí apenas entre os verbos de dois lugares os transitivos diretos e os transitivos oblíquos.

A vantagem de trabalhar com todas as ocorrências, incluindo repetições, num *corpus* de língua falada é poder contar com o princípio de que a não ocorrência possa corresponder à baixa frequência no uso em geral ou de baixa produtividade no sistema linguístico, ou o efeito dos dois fatores juntos. No entanto, a falta de ocorrência de nomes derivados de verbos bitransitivos deve ter mais a ver com baixa frequência de uso na amostra, já que, em português, é produtiva a derivação de nomes de verbos de três lugares, como *doação*, *envio*, *entrega* etc.

O principal critério empregado na análise da estrutura argumental das nominalizações é o da codificação dos participantes, um fator de análise que está, em geral, presente em estudos funcionalistas (Mackenzie, 1987; Koptjevskaja-Tamm, 1993; Cristofaro, 2003; Malchukov, 2004) que tratam das construções encaixadas, isto é, é relevante saber se os participantes do evento dependente são expressos como nas orações declarativas independentes; saber, por exemplo, se os participantes podem ou não ser claramente expressos, ou se são expressos como possuidores ou oblíquos. Como a nominalização, como estado de coisas que é, representa uma construção encaixada, a mesma metodologia será aplicada a ela.

Assim, nenhuma outra restrição foi operada no levantamento, a não ser a natureza valencial da nominalização em relação ao verbo subjacente correspondente, nem mesmo os processos formais de derivação.

Sabe-se que, do ponto de vista da morfologia derivacional, a formação de nomes a partir de verbos pode ser feita tanto por sufixação, como *construção*, quanto pelo processo de derivação regressiva, como *pesca*. Segundo Basílio (2004) as mais produtivas são as formações de estrutura $[\text{V-ção}]_N$, chegando a corresponder a cerca de 60% das formações regulares, seguidas por formações em –mento, que correspondem a cerca de 20% das formações regulares (Basílio, 2004, p.42). É provável que a amostra de nominalizações aqui analisada mantenha essa proporção, mas esse aspecto formal não é pertinente a este trabalho, que, como mencionado acima, é controlado pela natureza valencial da nominalização.

Os critérios para o levantamento da amostra não excluem a possibilidade de ocorrências repetidas do mesmo caso, dado que, mesmo estando muito próximas na superfície textual, duas repetições nem sempre exercem a mesma função discursiva. Compare, por exemplo, as ocorrências repetidas de *evolução* e *apresentação* nos fragmentos textuais contidos em (4-1) e (4-2) respectivamente.

(4-1) *no final das contas toda a <u>evolução</u> humana... não deixa de ser exatamente a <u>evolução</u> do domínio que o homem tem sobre a natureza...* (EF-SP-405)

(4-2) *Doc.:* *Dona I. além do filme em si a senhora acha que há alguma <u>apresentação</u> anterior que prende a atenção do público e até que ponto ela pode ser interessante ou não?*
Inf.: *<u>apresentação</u> anterior do filme?...* (DID-SP-234)

Em (4-1) a primeira menção é menos específica e, portanto, menos referencial que a segunda. Em (4-2), a primeira e a segunda menção de *apresentação* não têm a mesma referência, dado que o Documentador quer se referir a outro tipo de apresentação que

anteceda a do filme, não a do filme propriamente dita, enquanto a Informante considera, em seu pedido de esclarecimento, a possibilidade de terem sido mencionadas duas apresentações do mesmo filme na pergunta do Documentador.

É possível inferir que a amostra representa bem a produtividade do processo de derivação nominal na gramática do português, como se observa na Tabela 1, que traz uma distribuição dos casos cruzando os tipos de valência com os tipos semânticos de predicados.

A Tabela 1 faz menção ao predicado *input* das nominalizações. Assim, *V1* representa predicados monovalentes; *V2*, predicados bivalentes com complemento direto e *V2-Ob*, predicados bivalentes com complemento oblíquo, que, na tradição gramatical, são classificados como transitivos indiretos.

Tabela 1 – Valência potencial e tipos semânticos de predicados[2]

	V1		V2		V2-Ob		Total	
	N	%	N	%	N	%	N	%
Processo	24	41,0	04	4,0	2	6,0	30	17,0
Posição	1	1,0	7	8,0	16	50,0	24	13,0
Ação	24	41,0	66	73,0	1	3,0	91	50,0
Estado	10	17,0	13	15,0	13	41,0	36	20,0
Total	59	33,0	90	50,0	32	18,0	181	

A Tabela 1 mostra que os nomes derivados de verbos bivalentes com complemento direto são majoritários, seguidos dos derivados de verbos monovalentes. Registra-se uma baixa frequência relativa na amostra da classe de nomes derivados de verbos bivalentes com complemento oblíquo.

Os nomes de ação predominam sobre os demais tipos semânticos de estados de coisas, e predominam também na classe dos nomes mono, bi e trivalentes. A exceção fica por conta dos predi-

2 Nesta tabela e nas demais, N representa a frequência numérica bruta e %, a frequência percentual. *Valência potencial* representa aqui a estrutura argumental do verbo *input* em estado de dicionário.

cados nominais derivados de verbos com complemento oblíquo. Registra-se apenas um caso de estado de coisas de ação, frequência que é absolutamente compatível com a produtividade semântica desse tipo de predicado, que costuma ser muito maior, embora não absoluta, na classe dos predicados de estado e de processos com sujeito experienciador.

Uma característica secundária que merece observação é a função sintática que os nomes derivados exercem na predicação hierarquicamente superior. Há 10,0% (19/181) de casos de verbo suporte, como se observa em (4-3a-c).

(4-3) a *três ou quatro citações que <u>faziam referência</u> exatamente **a isso** que estilo mudava* (EF-SP-405)

b *como eu falei para vocês ele <u>faz seleção</u> **de pessoal** né?* (D2-SP-360)

c *o paleolítico é o período... da pedra lascada...como vocês todos sabem... não é?... e... <u>tem uma duração</u> **de aproximadamente seiscentos mil** anos* (EF-SP-405)

Quando complemento de verbo-suporte,[3] a nominalização não pode ser considerada propriamente um tipo de construção encaixada por atuar como segundo componente de uma espécie de lexia composta. Nesses casos, a nominalização tem o apoio de um verbo semanticamente mais leve – em geral *fazer, ter* e *dar* – que recebe todos os sufixos verbais. Isso não impede que, em alguns casos, ela apareça relativamente independente com marcação própria de determinação. Dizer, portanto, que *referência*, *seleção* e *duração* ocupam a posição de A[2] nessas predicações é um equívoco interpretativo, já que "examinando-se os pares formados em português pelas

3 Para Neves (2002), "uma definição corrente de verbo-suporte os apresenta como 'verbos semanticamente vazios que permitem construir um SN com V-n em relação de paráfrase com um SV [...]", é indicação que não deve ser tomada como definitória por haver construções desse tipo que não possuem correlatos semânticos com verbos simples (Neves, 2002, p.210).

construções com verbo-suporte, de um lado, e as construções com verbos plenos correspondentes, de outro, verifica-se que o contraste entre esses enunciados em relação parafrástica não implica diferenciação sintática no restante da oração" (Neves, 2002, p.211).

Os três casos das lexias acima mencionadas do tipo verbo+nominalização poderiam ser substituídos, sem prejuízo semântico, por verbos correspondentes, como *referiam-se*, em (4-3a), *seleciona*, em (4-3b), e *dura*, em (4-3c); são, portanto, autênticos complementos de verbo-suporte. Entretanto, se, por um lado, todos os casos dispõem de valência (o constituinte em negrito), por outro, as nominalizações de (4-3a) e (4-3b) são completamente destituídas de determinação, o que é típico de complemento de verbo-suporte, enquanto a de (4-3c) aparece acompanhada de artigo indefinido. De qualquer modo, observe que, nas ocorrências (4-3b-c), os nomes deverbais expressam a estrutura valencial.

Extraídas essas 19 ocorrências de nominalizações com verbo--suporte, a maioria absoluta, equivalente a 52,0% (84/162) dos casos, exerce função argumental correspondente à de sujeito (4-4a), de objeto (4-4b), de oblíquo (4-4c) do verbo *input* respectivo, e de complemento nominal (4-4d) na predicação matriz:

(4-4) a *as:: manifestações artísticas*[4] *começaram a aparecer no paleolítico superior...* (EF-SP-405)
b *eles também precisavam acompanhar... o a migração da caça se não eles iam fica sem comer...* (EF-SP-405)
c *e o problema de horários de adaptar a carreira... com... a a:: [ ] com a casa com a administração da casa* (D2-SP-360)
d *uma última coisa que eu gostaria de dizer é o fato de que nessa época ainda não existe preocupação com composição...* (EF-SP-405)

4 Outras referências anteriores justificam considerar que o adjetivo *artísticas* neste SN pode ser considerado como equivalente ao SP *de arte*, já que o principal tópico do texto é a arte do paleolítico.

Há também ocorrências de outras funções não majoritárias, tipicamente não argumentais, em que a nominalização atua como núcleo na função de satélite (4-5a), modificador de nome (4-5b), predicado não verbal (4-5c), e mesmo constituinte extraoracional (4-5d) (Dik, 1987), tratado na GDF como ato discursivo independente de Orientação.

(4-5) a *numa família grande há sempre um com tarefa de supervisor... por instinto não por obrigação...* (D2-SP-360)
b *houve uma série de irre/eh:: de irregularidades... nas lis/ na apresentação da lista de classificação* (D2-SP-360)
c *não e/ então não era divertimento aquilo...* (DID-SP-234)
d *quanto à coleta se eles dependiam... da colheita... de frutos... raízes... que eles NÃO plantavam...* (EF-SP-405)

Como todas as nominalizações são formas dependentes exercendo diferentes relações gramaticais na predicação matriz, os dados acima mostram a relevância dos deverbais como um mecanismo de subordinação. A causa dessa distribuição é que eles podem atuar não só como complemento nas funções argumentais e predicacionais da matriz, mas também como modificador nominal e adverbial dela e até mesmo de forma mais ou menos independente da matriz, na função de ato discursivo de Orientação, como em (4-5d).

Embora as outras ocorrências sejam minoritárias, o exercício de diferentes funções na predicação hierarquicamente superior, registrado nos dados acima, revela uma propriedade característica das nominalizações do português e de outras línguas, que é sua versatilidade funcional: como as construções nominais dispõem de outro tipo de complexidade categorial, elas oferecem menos resistência a restrições sintáticas (Mackenzie, 1985).

Observe-se particularmente o caso de (4-6) em que se registram encaixamentos recursivos em que *repressão* funciona como construção dependente, complementando o nome *campanha* na predicação matriz, que, por seu lado, é, ele próprio, matriz de *mendicância,* outro nome derivado.

(4-6) *fazendo parte da:: campanha de:: repressão à mendicância... do governo Carvalho Pinto* (D2-SP-360)

Além desse caso de recursividade, a versatilidade dos nomes derivados é ainda maior quando a predicação nominalizada pode exercer funções sintáticas na predicação matriz que não seriam licenciadas aos predicados primitivos correspondentes, como mostram os exemplos (4-7a-b) e (4-8a-b).

(4-7) a * *as artes se manifestarem [e isso começou a...] começaram a aparecer no paleolítico superior...*
b *as:: manifestações artísticas começaram a aparecer no paleolítico superior...* (EF-SP-405)

(4-8) a * *fazendo parte da:: campanha de:: que o governo Carvalho Pinto reprimiu a mendicância*
b *fazendo parte da:: campanha de:: repressão à mendicância... do governo Carvalho Pinto* (D2-SP-360)

Como se pode observar, a conversão das nominalizações *manifestação,* em (4-7b), e *repressão*, em (4-8b), na sua contraparte verbal, apresentada, respectivamente, em (4-7a) e (4-8a), resulta, nesses contextos específicos, em construções malformadas ou no mínimo estranhas. Esse fato está plenamente de acordo com a correlação entre tipos morfossintáticos de subordinação e grau de integração semântica entre predicação encaixada e predicação matriz.

A nominalização pode atuar como um exemplo de motivação funcional por iconicidade (Haiman, 1983; Givón, 1980), segundo o qual a integração semântica entre os eventos se reflete na integração morfossintática entre as orações, conforme a escala oração finita > oração não finita > nominalização (Lehmann, 1988). A nominalização consiste no grau máximo de integração formal no processo de gramaticalização por se aproximar em grau máximo de um SN atuando como complemento da oração matriz, conforme a gradação

sugerida por Mackenzie (1996), contida em (3-29a-d), repetida aqui como (4-8).

(4-8) a *That* ***my horse*** (A^1) *won* ***the race*** (A^2) *came as a great surprise*. (oração finita)

b ***My horse*** (A^1) *winning* ***the race*** (A^2) *came as a great surprise*. (gerúndio)

c ***My horse's*** (satélite-y) *winning* ***the race*** (A^2) *came as a great surprise*. (gerúndio-genitivo)

d ***My horse's*** (satélite-z) *winning* ***of the race*** (satélite-y) *came as a great surprise*. (nominalização produtiva)

De (4-8a) a (4-8d) observa-se uma gradação progressiva em direção ao grau máximo de dessentencialização, um dos parâmetros postulados por Lehmann (1988) para propor uma tipologia de ligação de orações, conforme disposto na Figura 7:

(Adaptado de Lehmann, 1988, p.189)

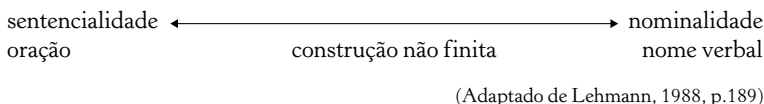



Figura 7 – Escala de dessentencialização

Baseando-se em Dik (1985) e, especialmente, Mackenzie (1985), Lehmann (1988) destaca que o processo de dessentencialização afeta a relação do predicado verbal com seus argumentos. Numa nominalização, os SNs correspondentes aos argumentos do verbo finito aparecem no genitivo ou em sintagmas adposicionais. Um verbo não completamente nominalizado pode ter um objeto direto e um advérbio; já um verbo mais nominalizado é modificado por um adjetivo, como um nome. Quanto mais a construção encaixada se aproxima do polo direito do *continuum*, mais o comportamento dela se aproxima ao de um membro prototípico da classe dos nomes. Nesse sentido, pode-se falar de um aumento progressivo no grau de nominalidade da oração subordinada, conforme sejam reduzidas pelo processo de dessentencialização.

4.2 Especificação e não especificação de valência

Considere-se, inicialmente, a Tabela 2 abaixo, que trata da relação entre valência potencial e expressão dos argumentos.

Tabela 2 – Relação entre valência potencial e expressão argumental

	V1		V2		V2-Ob		Total	
	N	%	N	%	N	%	N	%
A^1 expresso	21	84,0	4	16,0	/	/	25	14,0
A^2 expresso	/	/	33	92,0	3	8,0	36	20,0
A^1 e A^2 expressos	/	/	/	/	5	100,0	5	3,0
A^1 e A^2 não expressos	38	33,0	53	46,0	24	21,0	115	63,0
Total	59		90		32		181	

Observados os resultados gerais referentes ao total de nominalizações com a estrutura argumental apenas expressa em torno do núcleo do sintagma nominal (ver última coluna da direita), o tratamento quantitativo mostra que raramente se expressa a valência do verbo primitivo correspondente. Com efeito, os casos de expressão formal mostram que:

(i) apenas 14,0% (25/181) dos nomes têm apenas o primeiro argumento (A^1) expresso e apenas 20,0% (36/181) dos nomes têm somente o segundo argumento (A^2) formalmente expresso;[5]

(ii) há apenas 3,0% (5/181) de ocorrências com A^1 e A^2 expressos simultaneamente;

(iii) esses índices significam que há uma incidência de 63,0% (115/181) de nominalizações sem expressão de ambos os argumentos, o que constitui a esmagadora maioria dos casos.[6]

5 Como seria estranho mencionar funções sintáticas de sujeito e de objeto para a nominalização, o uso de A^1 e de A^2 representa mais apropriadamente funções semânticas de agente e paciente que nomes deverbais compartilham com o verbo por herança derivacional.

6 Ver também análises alternativas em Camacho e Santana (2004) e Santana (2005).

Apesar disso, argumentaremos, a seguir, com uma distribuição mais específica de dados, que o que ocorre com o português, e provavelmente com outras línguas, não é redução de valência, mas apenas expressão ou não do argumento subjacente por razões de natureza semântica e pragmática.

Considerando-se apenas a distribuição majoritária em relação ao total de argumentos expressos na última coluna da direita, observa-se que 84,0% (21/25) dos A^1 são expressos na estrutura argumental de nome derivado de verbo monovalente ou intransitivo. Considerando-se, por outro lado, o total de argumentos A^2 manifestos, observa-se que 92,0% deles (33/36) são expressos na estrutura argumental de nome derivado de verbo bivalente ou transitivo, e um conjunto pouco representativo de dados – todos os cinco casos de expressão simultânea de A^1 e A^2 – ocorre na estrutura argumental de um nome derivado de verbo transitivo oblíquo, o que seria de esperar, em virtude de A^2 herdar naturalmente a codificação gramatical de oblíquo do termo primitivo. Essa codificação é basicamente diferente da expressão com sintagma-de, codificação que incide somente sobre a nominalização para visibilizar os casos de sujeito e/ou de objeto do predicado verbal *input*.

Os dados sugerem que, no caso de estrutura bivalente, há uma forte competição em torno da expressão de um dos argumentos do nome, como prevê Dik (1985), e é o argumento paciente que vence. No caso de nomes derivados de intransitivos, em que não há competição, o argumento majoritariamente expresso é realmente o agente ou equivalente para A^1. Nos nomes derivados de verbos transitivos diretos, quando pelo menos A^2 é expresso, a competição se resolve principalmente pela oposição entre sintagma-de para a expressão de A^2 e anáfora zero para a expressão de A^1, como mostra o exemplo (4-9a), ou sintagma-de para A^2 e expressão indeterminada de A^1, como o exemplo (4-9b).

(4-9) a *eles precisam pegar pele para se esquentar... e ter comida... para comer e se defender dos outros animais... então as* <u>*preocupações*</u> *são MUITO... ahn::...* ***de todo (o) DIA****::* (EF-SP-405)

b *de acordo com o edital a validade é de dois anos DA publicação... **dos resultados**... da lista de aprovados...* (D2-SP-360)

Importante notar que os cinco casos de nominalizações que receberam manifestação simultânea de dois argumentos são verbos bivalentes com complemento oblíquo, como (4-10a) e (4-11a) e a forma verbal correspondente em (4-10b) e (4-11b) respectivamente.

(4-10) a *toda e qualquer manifestação que a gente for procurar vai ter que estar necessariamente ligada... a esta preocupação vital **do homem pré-histórico de... se conservar vivo**...* (EF-SP-405)

b *O homem pré-histórico se preocupa vitalmente com/em conservar-se vivo*

(4-11) a *Doc.: Dona I. além da participação **do artista**... **no filme** quais os outros elementos importantes na sua opinião para que o filme seja bem-sucedido bem aceito pelo público?* (DID-SP-234)

b *O artista participa do (no) filme*

Em (4-10) os dois argumentos da nominalização estão expressos como sintagma-de, situação que é geralmente evitada em português. No entanto, o falante altera a regência esperada do verbo, inserindo *de* em lugar de *com*, e o resultado não parece estranho porque o segundo argumento de *preocupação* não é um sintagma nominal, mas uma oração com verbo não finito.

Em suma: há uma baixa incidência de casos de nominalizações bivalentes com os dois argumentos expressos, como exemplificado em (4-10a); esses casos correspondem a predicados verbais *input* cuja estrutura subjacente já contém uma preposição, conforme ilustrado em (4-10b). Tais aspectos indicam claramente que a ausência de expressão formal dos argumentos é uma característica muito marcada dos dados, principalmente porque é um fenômeno geral que se aplica também aos predicados monovalentes.

Nota-se também algumas ocorrências de A^2 sem sintagma-de, como por exemplo (4-12), em que o uso de *sobre* em lugar de *de* pode ter sido semanticamente motivado.

(4-12) *que no final das contas toda a evolução* ***humana****... não deixa de ser exatamente a evolução do domínio* ***que o homem tem sobre a natureza****...* (EF-SP-405)

Prevalece a noção de posição superior que *sobre* sustenta, mesmo quando a noção concreta não mais se aplica, como se observa em (4-12), em que a seleção do nome *domínio* pode ter ativado no falante o conceito similar e mais abstrato de domínio, que não passa de uma abstração da noção espacial mais concreta.

Além disso, o mesmo exemplo apresenta uma ocorrência de nome deverbal em que a expressão de A^1 aparece sob a forma de adjetivo. Essa expressão formal certamente tem pouco a ver com a competição entre A^1 e A^2 por ocupar a posição de possuidor, mas exerce uma função textual extremamente relevante, que é retomar como adjetivo o SN *o homem (do paleolítico superior)*. Há um número reduzido de ocorrências desse tipo, que se manifestam majoritariamente no contexto de verbo intransitivo subjacente, como (4-12), e, em menor escala, no contexto de verbo transitivo, como em (4-13) e (4-14).

(4-13) *porque isso a gente vem dizendo até agora certo? se ela foi criada... para um FIM... OUtro... que NÃO... a contemplação* ***estética****...* (EF-SP-405)

(4-14) *ah::: por parte do público... depois de uma representação* ***teatral****?* (DID-SP-234)

Para nomes como *contemplação* e *representação*, a arena típica para competição argumental não ocorre com dois argumentos expressos, como nos exemplos de (4-10), já que, por indeterminação referencial do argumento A^1 do predicado transitivo de (4-13) e do argumento A^2 do predicado transitivo de (4-14), não há relevância pragmática em especificá-los formalmente. Além disso, a inserção de adjetivo para a representação de valência torna o predicado nominal deverbal mais próximo de um membro prototípico da classe dos nomes.

Outro caso possível é A^1 receber forma de pronome em posição pré-nominal e A^2, expressão pronominal em posição pós-nominal, que é a regularmente indicada para esse tipo de modificador. Assim a referência a possuidor não se restringe a sintagmas adposicionais. Os dados apontam para um único caso de expressão pronominal de A^1 e, até um pouco aparentemente estranho, para um caso de expressão de A^2, como mostram respectivamente os exemplos (4-15) e (4-16).

(4-15) *agora neste momento eu vou trabalhar com barro vou fazer* ***minhas*** *criações ou eu vou pintar um quadro...* (EF-SP-405)

(4-16) *enquanto na/ não for resolvido esse projeto não o projeto que tem... sabe? para os procuradores uma lei... nossa uma regulamentação* ***nossa*** (D2-SP-360)

É necessário observar que a referência pronominal em (4-15), tanto do pronome pessoal *eu* quanto de sua recuperação anafórica no contexto da sintaxe interna da nominalização como o possessivo *minhas*, tem uma referência genérica, implicando qualquer ser humano. De qualquer modo, este, que seria um caso típico de bloqueio do argumento sujeito por compartilhamento semântico, acabou recebendo expressão argumental justamente em virtude da diferença de manifestação formal, o que torna um tanto enfático, focal, o uso do possessivo no nome deverbal.

O mesmo se aplica ao exemplo (4-16), em que o possessivo *nossa* é também empregado num sentido genérico, mas implicando uma classe de pessoas, a de procuradores, que também inclui o falante. O nome *regulamentação* também nesse contexto seria desprovido de expressão argumental, mas a expressão pronominal que recebe não é redundante, já que o possessivo acrescenta uma informação relevante por mostrar que o falante se inclui na categoria dos procuradores.

Uma explicação possível para a alta taxa de não especificação formal é que não há relevância discursiva para expressar valência, que é, em geral, suprida pela organização do discurso. Se já há in-

formação contextual e cotextual suficiente sobre os argumentos, é descabido expressá-los na estrutura interna do predicado nominal. Nesse caso, a ausência de expressão formal não significa redução argumental no processo de derivação do verbo para o nome; ao contrário, significa competição entre expressão por anáfora zero e por sintagma adposicional.

4.3 A codificação morfossintática dos argumentos

Para iniciar a discussão, vamos observar a distribuição dos argumentos da nominalização por tipo de expressão formal, conforme exposto na Tabela 3.

Considerados os totais, a frequência de expressão de A^2 (34,0%) é mais que o dobro da expressão de A^1 (16,0%), o que é de fato uma distribuição esperada com base nas predições de Dik (1985, 1997).

Tabela 3 – Forma de expressão dos argumentos

Expressão argumental ➔	A^1 expresso		A^1 não expresso		A^2 expresso		A^2 não expresso	
Forma de expressão:	N	%	N	%	N	%	N	%
Sintagma-de	22	12,0			30	25,0		
Adjetivo	4	2,0			1	1,0		
Outra expressão	4	2,0			10	8,0		
Total	**30**	**16,0**			**41**	**34,0**		
SN equivalentes			44	24,0			6	5,0
Anáfora zero			47	26,0			32	26,0
Argumento no texto seguinte			4	2,0			4	3,0
Argumento indeterminado			26	15,0			17	14,0
Predicado não referencial			30	17,0			22	18,0
			151	**84,0**			**81**	**66,0**
Total			**181**				**122**	

Vale a pena lembrar que, se por um lado ambos os argumentos recebem a expressão preferencial de sintagma-de, essa forma atribuída a A^1 só é a preferencial para os casos monovalentes; havendo dois argumentos disponíveis na estrutura argumental, predomina a expressão de A^2 como sintagma-de, e A^1 recebe, nesse caso, outras formas de expressão, como adjetivo ou pronome possessivo. Esses resultados estão plenamente de acordo com as regras de expressão para as nominalizações postuladas por Dik (1985) e já discutidas na Seção 3.2.

Além disso, os dados confirmam a predição de que forma de expressão argumental preferida é a de sintagma adposicional para a posição de possuidor, já que, do total de A^1 expressos, 74,0% (22/30) e, do total de A^2 expressos, 73,0% (30/41) aparecem sob essa forma de manifestação numa relação quantitativa de surpreendente simetria. Os exemplos (4-17) e (4-18) mostram casos de A^1 e A^2 respectivamente expressos como sintagma-de como se fosse expressão de possuidor.

(4-17) L2 *e isso significa um aumento* ***de vencimentos****...* (D2-SP-360)

(4-18) L2: *tudo... de a responsabilidade na manutenção* ***da casa****...*
L1: *vem como complemento né?* (D2-SP-360)

Considerando-se agora a forma de manifestação de A^1 e A^2, nota-se que, na coluna "A^1 expresso" (significando manifestação em torno do próprio predicado nominal), há um total de trinta casos na Tabela 3. Na coluna "A^1 não expresso" (significando argumento não expresso em torno do predicado nominal), há 44 casos de expressão por anáfora zero de argumentos semanticamente compartilhados com o predicado da matriz, conforme demonstra o exemplo contido em (4-19a-b).

(4-19) a *ajudar um pessoal que que que tem me pedido para fazer:: programação da da de sucos do Lanjal e eu acho que a televisão é completamente:: diferente do que a gente assiste eh lá no no teatro não o teatro::...* (DID-SP-234)

b ** ajudar um pessoal que que que tem me pedido para fazer:: programação de sucos do Lanjal (por mim)*

O argumento agentivo de *programação* é zero catafórico, já que sua referência remete para o complemento dativo do predicado encaixador; nesse caso, a inserção de um termo na posição argumental, como em (4-19b), torna a construção agramatical por razões de redundância.

Há também 47 casos de expressão por anáfora zero cujo argumento não é compartilhado com a predicação encaixadora, mas que retomam anaforicamente algum outro SN contido no contexto discursivo, conforme aparece em (4-20a-b).

(4-20) a *então numa vida desse tipo... a preocupação principal está centrada na sobrevivência...* (EF-SP-405)

b *então numa vida desse tipo... a preocupação principal do homem está centrada na sobrevivência...* (EF-SP-405)

Por fim, a amostra registra, além disso, mais quatro casos de argumentos não expressos na estrutura do SN nucleado pelo nome deverbal, mas que são recuperáveis no contexto seguinte, conforme especificado em (4-21a-b).

(4-21) a *é MUIto difícil a gente desenhar estritamente o que a gente vê a gente separar a percepção... da... do conceito que nós fazemos do objeto...* (EF-SP-405)

b ** separar a percepção do objeto* ***por nós*** *do conceito que* ***nós*** *fazemos do objeto*

O argumento experienciador de *percepção* é zero catafórico, já que sua referência remete para o sujeito de uma oração relativa, atuando como modificador do núcleo nominal de *conceito*, que aparece, por seu lado, sob a forma de sintagma-de na condição de segundo argumento do nome deverbal. Como (4-19a-b), a inserção de um termo na posição argumental torna a construção também inaceitável por razões de redundância.

Somados esses três tipos de manifestação argumental com os de expressão formal no interior do SN, há, na realidade, uma incidência de 68,0% (125/181) de manifestação de A^1 em oposição a 32,0% (56/181) de predicados sem expressão referencial (trinta casos) e de argumento sujeito semanticamente indeterminado (26 casos). O exemplo contido em (4-22) ilustra um caso de nominalização cujo argumento A^1 é semanticamente indeterminado.

(4-22) *e e:: além de que... da/dentro do aumento de vencimentos haveria... uma promoção de todo o pessoal que está agora...* (D2-SP-360)

A promoção a que o falante se refere em (4-22) deve ser realizada pelas autoridades do Ministério da Justiça que regem a carreira de Procurador do Estado, mas ele/ela não vê relevância discursiva na menção do agente, exatamente como ocorre em construções passivas do português (*o pessoal foi promovido /promoveu-se o pessoal*) ou outras construções com predicados verbais com sujeito indeterminado (*promoveram o pessoal*).

Os casos contidos em (4-23a-c) são todos exemplos de predicados não referenciais.

(4-23) a *Então (para a) mulher aquilo é um complemento quer dizer tudo que vem é ótimo* (D2-SP-360)

b *eu hav::ia lido um:: sobre um novo método de alfabetização...* (D2-SP-360)

c *a senhora falou em divulgação que:: talvez o teatro é menos aceito pelo público por falta de divulgação certo?* (DID-SP-234)

As ocorrências de (4-23a-c) constituem casos de nomes prototipicamente não referenciais, como os que ocorrem como predicado não verbal em (4-23a), os que ocorrem como SPs em função de modificador como em (4-23b), ou como complemento de nome como em (4-23c).

Nesses casos, o nome, em si mesmo designando uma propriedade (f_1), fornece apenas a descrição lexical de um estado de coisas. O nome se aproxima do estatuto de uma categoria de nível zero, que devem ser consideradas as menos prototípicas das categorias referenciais: como o referente de uma expressão de nível zero, essas nominalizações não podem realmente ser consideradas uma entidade, mas apenas a propriedade denotada por elas (cf. Keizer, 2004b).

Em casos como esses de suspensão de referencialidade prototípica, nomes de ordem superior, como as nominalizações, seguem a mesma tendência do nome prototípico de primeira ordem na mesma situação discursiva, isto é, a não referencialidade se manifesta na ausência de determinantes, modificadores, quantificadores, configuração formal típica de nominais nus. Vale lembrar que, na escala de predicatividade proposta por Hengeveld (1992) – predicado nu > predicado referencial > predicado relacional –, o uso de nominais nus é a estratégia mais usada para representar predicado não verbal. Nesse caso, não se pode falar em redução de valência, já que fica suspensa a referência a uma propriedade ou relação, própria de entidades de ordem superior.

Aplicado o mesmo raciocínio sobre os argumentos A^2, os resultados são surpreendentemente comparáveis: somados os casos de A^2 expressos (n=41), os de argumentos não expressos por serem semanticamente compartilhados (n=6), os de argumentos pragmaticamente retomados por anáfora zero (n=26), e os que aparecem no contexto textual seguinte (n=4), obtém-se uma frequência de 68,0% (77/122) de manifestação argumental, em oposição a 18,0% (22/122) de argumentos A^2 indeterminados e 14,0% (n=17) de predicados não referenciais. Veja em (4-24a-b) uma ocorrência de argumento semanticamente compartilhado entre a predicação matriz e a predicação encaixada.

(4-24) a *e eles conseguem chegar... a é óbvio uma evolução certo?* (EF-SP-405)

b * *e eles conseguem chegar... a é óbvio uma evolução* ***deles*** *certo?* (EF-SP-405)

A não relevância da expressão da estrutura argumental do nome derivado tem a ver, portanto, com um dos princípios de cooperação postulados por Grice (1975), a Máxima da Quantidade, segundo o qual não se deve dizer mais do que o necessário e é possível aplicá-la com resultados satisfatórios às três modalidades de não especificação argumental. Esse princípio pragmático tem acolhida na GDF, que leva em consideração a responsabilidade dos participantes da interação pelo conteúdo comunicado.

4.4 Motivações semânticas e pragmáticas para a não especificação dos argumentos

Os dados discutidos na seção anterior mostraram que há casos diferentes de zero anafórico e de zero catafórico. Na literatura funcionalista essas motivações, vale destacar, estão ligadas ao princípio de economia, em contraposição ao de iconicidade (Givón, 1980, 1990; Haiman, 1983), duas motivações funcionais extremamente reconhecidas pela pesquisa tipológica.

A iconicidade é a tendência para a relação de conformidade entre a estrutura linguística e a estrutura da experiência conceitual, traduzida por Haiman (1985, p.11) como motivação icônica. A motivação icônica é uma correspondência entre a estrutura linguística e o significado tal que relações entre partes da estrutura linguística refletem diagramaticamente relações entre os conceitos que ela codifica.

Um dos domínios da motivação icônica é a chamada iconicidade de distância (Haiman 1985), ilustrada pela expressão de posse: a distância entre morfemas sinalizando posse inalienável nunca é maior que a distância entre morfemas sinalizando posse alienável. Outro domínio é fornecido pelo Princípio de Ligação (Binding Principle) de Givón (1980), segundo o qual o grau de integração semântica entre os eventos se reflete no grau de integração morfossintática entre a oração matriz e a encaixada (Givón, 1980; 1990).

Em contraste com a iconidade, que torna a estrutura linguística tão transparente quanto possível, a economia representa uma pressão para o mínimo esforço e para a simplificação máxima da expressão. Haiman (1985) faz uma distinção entre economia sintagmática e economia paradigmática. A economia paradigmática se traduz pela tendência para reduzir ao máximo o inventário de signos de um sistema linguístico, cuja consequência mais direta é o fenômeno da polissemia. A economia sintagmática, por outro lado, é a tendência para reduzir o comprimento ou a complexidade de qualquer enunciado, de modo que as expressões mais frequentes no uso tendem a reduzir-se fonologicamente e a informação que é redundante e/ou recuperável no contexto comunicativo tende a ser omitida (Haiman, 1985, p.158).

Segundo Cristofaro (2003, p.248-9), é a economia sintagmática que explica a correlação entre a predeterminação de traços semânticos entre dois estados de coisas conectados e os fenômenos morfossintáticos que levam à não especificação de informação correspondente, a que ela denomina Princípio de Recuperabilidade de Informação.

Em seus próprios exemplos, a economia sintagmática está fortemente correlacionada com a predeterminação de informação semântica em estados de coisas conectados e a não manifestação de informação pragmaticamente compartilhada pelos participantes da interação. Se, por exemplo, o tempo, o aspecto e o modo de uma predicação dependente são predeterminados pelos traços semânticos da predicação principal, eles prescidem de especificação no verbo da predicação dependente, conforme se pode ver em (4-25), que consiste num exemplo de construção complexa em Maricopa.

(4-25) Maricopa (Northern Amerindian, Hokan)

[nyaa	**m-mii-k**]	*´ii-m*
I	2-cry-SUB	say-REAL

'I said you cried'

(Gordon,1986, p.247 apud Cristofaro, 2003, p.66)

É óbvio que esse tipo de não especificação se aplica a nominalizações. Similarmente se duas predicações compartilham os mesmos participantes, a referência a eles também pode ser omitida na predicação dependente, como ilustrado em (4-26), que tem como predicado matriz um verbo volitivo.

(4-26) Guugu Yimidhirr (Australian)

Ngayu	*wawu-dgirr*	*[mayi*	*buda-nhu]*
1SG.NOM	want-COM.ABS	food.ABS	eat-PURPV

'I want to eat food'

(Haviland, 1979, p.135 apud Cristofaro, 2003, p.78)

Esse princípio, também mencionado por Koptjevskaja-Tamm (1993), explica casos de compartilhamento argumental em nominalizações como o contido em (4-27) cujo argumento A[1], *caça*, não pode ser especificado na estrutura valencial de *resistência* por motivo de redundância.

(4-27) *isto é de caça... que é o que oferece... uma resistência porque a:: fruta está lá então eles não precisavam se preocupar...* (EF-SP-450)

O resultado é, segundo Cristofaro (2003), tornar a estrutura linguística o mais opaca possível, em contraposição ao princípio da motivação icônica. Trata-se aqui, de acordo com Haiman (1983), do princípio de economia sintagmática ou discursiva que explica a tendência pela omissão de informação redundante ou recuperável no contexto.

Parece não destituído de razão considerar a ausência de representação argumental nos nomes não referenciais, expressos como nominais nus, e a ausência de argumentos sujeito e objeto genéricos também como casos funcionalmente motivados por economia (Haiman, 1983) e explicáveis pragmaticamente pela Máxima da Quantidade de Grice (1975).

Se a valência potencial pode ser expressa no exterior do núcleo nominal, é possível considerar como argumentos alguns tipos de termos não fonologicamente manifestos na posição de primeiro e de segundo argumento da nominalização, que são expressos por dois diferentes tipos de expressão por zero anafórico.

O primeiro caso de zero anafórico representa um participante, o primeiro argumento, que é semanticamente compartilhado com o predicado da oração matriz, como mostra (4-28a-b). Note-se, de passagem, que (4-28) traz especificação, como SN pleno, apenas do segundo argumento, que aparece como sintagma-de.

(4-28) a *se eles dependiam... da colheita... de frutos... raízes... que eles não plantavam* (EF-SP-405)
b *eles colhiam frutos, raízes*

Note que a mesma nominalização com o primeiro argumento especificado em (4-28c) não seria uma construção aceitável por razões de redundância, já que o agente já se acha mencionado no sujeito de *dependiam*, que funciona como núcleo da predicação matriz.

(4-28) c ** se **eles** dependiam... da colheita... **por eles** de frutos... raízes... que eles não plantavam*

Vale lembrar que essa condição se sustenta apenas se houver identidade de participantes entre dois diferentes estados de coisas. No caso de não identidade, como se vê em (4-28d), a gramática licencia o agente do nome deverbal.

(4-28) d *as mulheres dependiam... da colheita... pelos homens de frutos... raízes.... que elas não plantavam*

O segundo tipo de zero, que é também anafórico, representa termos que recuperam alguma entidade dada, já mencionada no contexto precedente, não necessariamente na predicação matriz, como se vê em (4-29a-b).

(4-29) a *nessa época ainda não existe preocupação com composição...* (EF-SP-405)

b *Nessa época o homem pré-histórico não se preocupa com composição.*

Como o tópico do texto é a arte do paleolítico, há várias menções anteriores ao homem pré-histórico. É por isso que, em (4-29c), a mesma nominalização com o argumento especificado continuaria sendo uma construção aceitável. O outro argumento é *composição*, que aparece na forma de oblíquo do predicado verbal *input*.

(4-29) c *nessa época ainda não existe preocupação do* ***homem pré-histórico*** *com composição...*

Isso significa que, nesse aspecto, (4-29a) é diferente de (4-28a): a manifestação ou não do primeiro argumento em (4-29a) é uma escolha real do falante, diferentemente de (4-28a), que bloqueia a manifestação do argumento na função semântica de agente.

Passemos, agora, ao exame de casos de zero anafórico na expressão do segundo argumento, conforme se vê no exemplo contido em (4-30a-b).

(4-30) a *ele percebeu que era capaz de CRIAR::... e criar uma imagem... [...] então:: ele vai tentar usar esta criação... que ele é capaz de fazer... para garantir a caça...* (EF-SP-405:52-3)

b *O homem pré-histórico criou uma imagem.*

O zero anafórico de (4-30a) representa um participante semanticamente compartilhado com o predicado da oração matriz na posição de segundo argumento, o que equivaleria, na predicação verbal de (4-30b), ao objeto ou argumento interno de *criar*.

A mesma nominalização com o segundo argumento especificado em (4-30c) é uma construção discursivamente menos aceitável por razões de redundância, uma vez que o paciente já se acha mencionado no objeto de *criar*, que aparece negritado.

(4-30) c *ele percebeu que era capaz de CRIAR::... e criar uma* ***ima-gem****... [...] então:: ele vai tentar usar esta* *criação* ***da imagem****... para garantir a caça...*

Os argumentos de *criação,* em (4-30a-b), que recebem ambos expressão de zero anafórico na sintaxe interna do predicado nominal, são facilmente recuperáveis no cotexto: o primeiro argumento é *ele* (homem pré-histórico) e o segundo é *imagem*; no entanto, o princípio de economia bloqueia o licenciamento da expressão sintática dos argumentos na retomada seguinte. Observe-se, todavia, (4-31) abaixo:

(4-31) *porque na medida... em que acabava a caça do lugar [...] eles também precisavam acompanhar... o a* *migração* ***da caça*** (EF-SP-405:49)

O argumento *caça* no SN grifado, que é o sujeito na predicação *input* correspondente, não é informação acessória, que possa ser descartada, uma vez que estabelece uma distinção desambiguadora no paralelismo entre migração do homem e migração da caça, sendo, nesse caso, imprescindível para a continuidade temática, para a coerência textual. Além disso, constitui informação pragmaticamente nova ou focal, em oposição à informação anterior, dada ou tópica, contida no SN *a caça do lugar*; a expressão sintática desse argumento tem, portanto, motivação semântica e pragmática.

Há um segundo tipo de expressão por zero, que não é anafórico, mas catafórico. Num número reduzido de casos, o argumento não se manifesta na sintaxe interna da nominalização por aparecer inserido no contexto imediatamente seguinte. Observe o exemplo contido em (4-32a-b) para ilustrar um caso de primeiro argumento.

(4-32) a *é MUIto difícil a gente desenhar estritamente o que a gente vê a gente separar a percepção... da... do conceito que nós fazemos do objeto...* (EF-SP-405:56)

b *Nós percebemos o conceito do objeto.*

O argumento experienciador de *percepção* é zero catafórico, já que sua referência remete para o sujeito de uma oração relativa, atuando como modificador do núcleo nominal de *conceito*, que aparece, por seu lado, sob a forma de sintagma-de na condição de segundo argumento do nome deverbal. Como (4-28a-b), a inserção de um termo na posição argumental torna a construção inaceitável também por razões de redundância, como se vê em (4-32c).

(4-32) c *separar a percepção do objeto* ***[*por nós]*** *do conceito que* ***nós*** *fazemos do objeto.*

O mesmo é aplicável ao segundo argumento, que também não se manifesta no interior da predicação matriz porque aparece inserido numa construção relativa imediatamente seguinte, conforme se observa em (4-33a-b).

(4-33) a *é inCRÍvel o que aparece lá os* <u>*cortes*</u> *que eles dão nas cenas* (DID-DP-234)

b *Eles cortam as cenas.*

Em (4-33) a referência aos argumentos do nome *cortes* está na oração relativa seguinte, o que corresponderia à representação subjacente contida em (4-33c).

(4-33) c $(e_i\ [f_i\text{: cortes}_N\ (f_i)\ (x_i\text{: eles}\ (x_i)_{Ag}\ (x_j\text{: cenas}\ (x_j)_{Pac})]\ (e_i))$

O argumento paciente de *corte* é zero catafórico, já que sua referência remete para o objeto oblíquo da oração relativa, atuando como modificador do núcleo nominal de *cortes*. A inserção de um termo na posição argumental torna a construção inaceitável também por razões de redundância, como se vê em (4-33d).

(4-33) d ** é inCRÍvel o que aparece lá os* <u>*cortes*</u> ***das cenas*** *que eles dão* ***nas cenas*** *é::*

Quando os participantes são predeterminados, o falante pode excluir a referência a eles porque a informação correspondente está implicada na relação de subordinação, ou em outro tipo de combinação de orações. Cristofaro (2003) não menciona os casos de compartilhamento argumental, como os tratados em (4-32a), que devem ser obviamente considerados como semanticamente motivados, já que a manifestação argumental por zero é cognitivamente compensada, por remissão catafórica, com um SN pleno ou pronominal na oração presente na oração relativa seguinte.

Há, no entanto, casos de compartilhamento de participantes em que a não especificação do argumento não é semanticamente motivada. Quando os participantes não são semanticamente predeterminados, e não há referência aberta a eles, a recuperação da informação depende do conhecimento de curto prazo pragmaticamente compartilhado pelos interlocutores. Um bom exemplo desse tipo de expressão argumental é (4-34a-b).

(4-34) a *toda e qualquer manifestação que a gente for procurar vai ter que estar necessariamente ligada... a esta preocupação vital do homem pré-histórico de... se conservar vivo [...] então a preocupação central... vai ser em torno da caça...* (EF-SP-405:50-1)

b *O homem pré-histórico se preocupa com conservar-se vivo.*

O exemplo (4-34a) contém duas menções do nome deverbal *preocupação* enunciadas em momentos diferentes do texto. Na primeira menção, os dois argumentos do nome deverbal estão manifestos: o primeiro argumento, que corresponderia ao sujeito do verbo *input*, é *o homem pré-histórico*, enquanto o segundo argumento, que corresponderia a um complemente oblíquo, é oracional (*de se conservar vivo*).

Já na segunda menção, fica claramente explícito que a expressão do primeiro argumento é igualmente o SN *o homem pré-histórico*, que, embora seja morfossintaticamente licenciado, é expresso por zero anafórico por consistir em informação pragmaticamente compartilhada pelos participantes da interação. Nesse caso, a segunda

menção da nominalização de (4-34a), aparentemente desprovida de manifestação valencial, não envolve perda, porque o ouvinte recupera uma informação de curto prazo, mencionada um momento antes na situação discursiva.

Segundo Cristofaro (2003, p.250-1), embora esse tipo de construção seja apenas outro exemplo do mesmo princípio de economia, ele reflete o fato de que essa situação de compartilhamento é mais comum no nível do discurso, e a organização estrutural é motivada por razões pragmáticas.

Em suma: os dados analisados mostram que há duas motivações para os diferentes tipos de zeros na expressão tanto do primeiro quanto do segundo argumento: (i) uma motivação semântica, que consiste nos casos de identidade de participantes mediante relação anafórica e catafórica; (ii) uma motivação pragmática, caso em que a identidade existe, mas a primeira menção está textualmente distante e pode ser recuperada mediante segunda menção.

Parte III

Generalizações e implicações teóricas

5
Uma interpretação funcional das construções nominalizadas

Introdução

Tendo comprovado que a estrutura valencial é propriedade constitutiva do predicado nominal derivado no léxico, volto, agora, a atenção, neste capítulo, para a tentativa de fornecer uma interpretação adequada do modo como o Nível Interpessoal e o Representacional estabelecem forte correlação com a forma de expressão da estrutura valencial no nível Morfossintático mediante o uso da nova arquitetura para a GF constituída pela GDF.

Vale a pena lembrar que tanto o predicado da matriz quanto o da oração encaixada numa relação de subordinação são marcados por diferentes graus de interconexão cujos efeitos são o compartilhamento de participantes e das marcas de tempo/aspecto/modo. Esse tipo de relação é conhecido na pesquisa tipológica como integração semântica (Givón, 1980, 1990; Noonan, 1985).

Neste capítulo, pretendo mostrar como encaixar esses fenômenos no âmbito da GDF mediante a interpretação do papel das nominalizações como construções encaixadas. Para tanto procedo conforme o seguinte roteiro.

Na Seção 5.1, que trata da implementação dinâmica, mostro que a posição assumida na análise é compatível com o Princípio de

Adequação Psicológica da GDF; na Seção 5.2, dando continuidade a esse princípio, proponho uma releitura para a tipologia dos predicados encaixados postulados por Dik (1997).

Na Seção 5.3, desenvolvo o princípio de que a separação entre predicados e esquemas de predicação, proposta no âmbito da GDF, é adequada para a interpretação de que zeros anafóricos são motivados por fatores semânticos e pragmáticos e, portanto, definidos nos níveis mais altos de organização gramatical.

Finalmente, na Seção 5.4, que fecha esse ciclo de análise com consequências teóricas, elaboro uma proposta de formalização para as nominalizações conforme o arcabouço da GDF.

5.1 A implementação dinâmica dos zeros anafóricos

Como visto na Seção 4.4, a economia sintagmática representa um princípio orientado pela tendência do falante para o mínimo esforço e a simplificação máxima da expressão, um princípio que explica a redução do comprimento ou da complexidade do enunciado com subsequente omissão de informação redundante ou recuperável no contexto comunicativo (cf. Haiman, 1985).

Segundo Cristofaro (2003) é a economia sintagmática que explica a correlação entre predeterminação de traços semânticos, como referência temporal, aspectual e modal na predicação matriz que conduz à não especificação da informação correspondente na predicação encaixada. Similarmente, se a predicação matriz e a encaixada compartilham participantes, também a referência a esses participantes pode ser omitida na forma dependente. Esse princípio, também mencionado antes por Koptjevskaja-Tamm (1993), explica casos de compartilhamento argumental como (4-26) aqui repetido como (5-1).

(5-1) *eu tenho impressão que* ***o teatro*** *PERde público por falta de* <u>*divulgação*</u> *[do teatro]...* (DID-SP-234:116)

Há, no entanto, casos de compartilhamento argumental em que o argumento não expresso não é semanticamente determinado pelo predicado da oração matriz. Quando os participantes são predeterminados, o falante pode excluir a referência a eles porque a informação correspondente está acarretada pela relação de subordinação. Todavia, quando os participantes não são predeterminados, e não há referência aberta a eles, poderia não haver modo algum de recuperar a informação perdida. O caso contido em (5-2b) é um bom exemplo desse tipo de não expressão argumental.

(5-2) a *toda e qualquer manifestação que a gente for procurar vai ter que estar necessariamente ligada... a esta preocupação vital do homem pré-histórico de... se conservar vivo...* (EF-SP-405:50)

b *então a preocupação central... vai ser em torno da caça...* (EF-SP-405:51)

Está claro em (5-2b) que a expressão de A^1 é o homem pré-histórico, enunciado alguns momentos antes no contexto, como aparece em (5-2a). O mesmo argumento na segunda menção está morfossintaticamente licenciado, mas não é expresso por ser informação pragmaticamente compartilhada entre os participantes da interação. Nesse caso, a construção de (5-2b) sem expressão fonológica de argumento não envolve perda real de informação, porque o ouvinte sabe que, se a construção de (5-2b) é assim usada pelo falante, é porque ela compartilha o participante experienciador não formalmente com o predicado matriz, mas pragmaticamente, na suposição de que a memória de curto prazo do ouvinte preserva a informação contida em (5-2a), que foi mencionada apenas um momento antes na situação discursiva.

Segundo Cristofaro (2003), embora esse tipo de construção seja apenas outro exemplo do mesmo Princípio de Recuperabilidade de Informação, ele reflete o fato de que essa situação de compartilhamento é mais comum no nível do discurso, e a organização estrutural está alinhada às condições do uso.

O Princípio de Recuperabilidade de Informação conforma-se adequadamente com o modo de organização da GDF. Os Níveis Interpessoal e Representacional alimentam-se igualmente de informações disponíveis no Componente Contextual, habilitando referência subsequente a várias entidades relevantes em cada nível, na medida em que elas são introduzidas no discurso. A operação de formulação apoia-se nesse componente de modo que a acessibilidade de antecedentes pode influenciar a composição dos atos discursivos em relação à visibilidade ou não de referentes no processo de codificação (Hengeveld; Mackenzie, 2006).

Há, portanto, uma forte interação entre os Níveis Representacional e Interpessoal nesse processo de codificação dos participantes na estrutura argumental do nome, justamente em razão de cada nível de representação dentro da gramática consultar o Componente Contextual, conforme mencionado acima. O falante recorre assim à memória de curto prazo do ouvinte, representada pelo componente contextual, para omitir a expressão fonológica do argumento de *preocupação* em (5-2b).

Nos casos de predicados não referenciais, como (5-3), e argumentos com referência indeterminada, como o argumento agente de (5-4), as condições semânticas, no caso do primeiro, e as pragmáticas, no caso do segundo, consistem em instruções para o ouvinte não esperar acessibilidade a nenhum referente específico, fenômeno que não é restrito às nominalizações.

(5-3) *ela quer saber as matérias que ela vai ter... o curso::... o segundo ciclo que ela pretende fazer... sabe? bom já está numa:: idade de definição quanto ao segundo ciclo porque elas já estão na oitava série as mais velhas não é?* (D2-SP-360)

(5-4) *aumenta um pouco mais a procura de engenheiro civil... depois cai...* (D2-SP-360)

Como ficou claramente demonstrado com base nos casos de anáfora zero, motivados por compartilhamento de participantes e

nos casos de predicados não referenciais, o Nível Representacional pode exercer um importante papel na etapa da formulação sem qualquer correlação evidente com o Nível Interpessoal, que é o mais alto da escala no processo de formulação.

Por isso, entende Mackenzie (2004a) que a hipótese de redução de valência licencia distintos esquemas para as categorias intermediárias já no Nível Representacional. Um dos argumentos mais fortes de Mackenzie é que as escolhas não são dependentes do Nível Interpessoal, já que, como se sabe, restrições de seleção determinadas pelo predicado encaixador nem sempre licenciam todas as formas alternativas na posição de construção dependente. Confira os exemplos que fornece, aqui sob o número (5-5a-b):

(5-5) a *I knew/said my horse won the race.*
b **I knew/said my horse ('s) winning (of) the race.*

(Mackenzie, 2004a)

De fato, os estados de coisas principal e dependente numa relação de subordinação podem ter vários graus de interconexão, em termos de compartilhamento de marcação de tempo, aspecto modo e também de participantes, fenômeno conhecido por integração semântica (Givón, 1980, 1990; Noonan, 1985). Com base nesse parâmetro, Cristofaro (2003) propõe uma relação entre predicados encaixadores e integração semântica, que reproduzo na Figura 8.

Integração semântica: Hierarquia de Integração Semântica	Ausência de Integração Semântica
Fasais > Modais > Manipulativos > ('fazer') > Manipulativos ('ordenar'), Desiderativos, Perceptivos	Conhecimento, Atitude Proposicional, Enunciativos

'>' = envolve maior integração semântica que'

(Cristofaro, 2003, p.122)

Figura 8 – Predicados encaixadores e integração semântica

Hengeveld (1998) faz referência ao fato de que, no caso da complementação, a natureza da representação formal do complemento oracional está fortemente correlacionada com a natureza semântica do predicado encaixador. Assim, verbos que descrevem atos de fala, como *say*, tomam uma oração subordinada como complemento; verbos expressando atitude proposicional, como *believe*, uma proposição subordinada como complemento; e finalmente verbos como *bother*, que expressam o efeito de um estado de coisas sobre um experienciador, tomam uma predicação subordinada como complemento. Essa dependência permite fazer fortes predições em relação à possibilidade de exprimir operadores de TAM: quanto mais alta a camada mais externa de uma oração de complemento, mais provavelmente ela será expressa por uma forma verbal independente.

Essas evidências de integração semântica, abundantemente fornecidas na literatura, mas postuladas especialmente por Givón (1980), mostram que a expressão de TAM e de participantes age, na formulação, diretamente sobre o Nível Representacional, antes de qualquer operação significativa com o Nível Interpessoal; nesse caso é ainda no próprio Nível Representacional que se deve buscar explicação para os casos de anáfora zero que, se consistem numa escolha do falante, essa escolha é feita indiretamente mediante mecanismos de economia sintagmática: a enunciação de uma sentença complexa com compartilhamento de pelo menos um participante implica morfossintaticamente a necessidade de eliminar a manifestação do argumento na nominalização encaixada que já estava especificado na oração matriz (ver Seção 4.4).

Entretanto, que explicação dar aos casos de anáfora zero motivados por seleção do falante em função da acessibilidade presumida de referente na memória de curto prazo do ouvinte? A resposta mais plausível para essa questão, em defesa da hipótese de preservação de valência, reside na interação dinâmica entre os níveis.

Antes de dar uma interpretação a esses casos, vale a pena lembrar que a GDF consiste num modelo hierárquico de gramática, constituído por um sistema multifuncional de níveis e camadas,

conforme demonstrado na Figura 3. O Nível Interpessoal e o Representacional operam na formulação, e o Nível Morfossintático e o Fonológico operam na codificação.

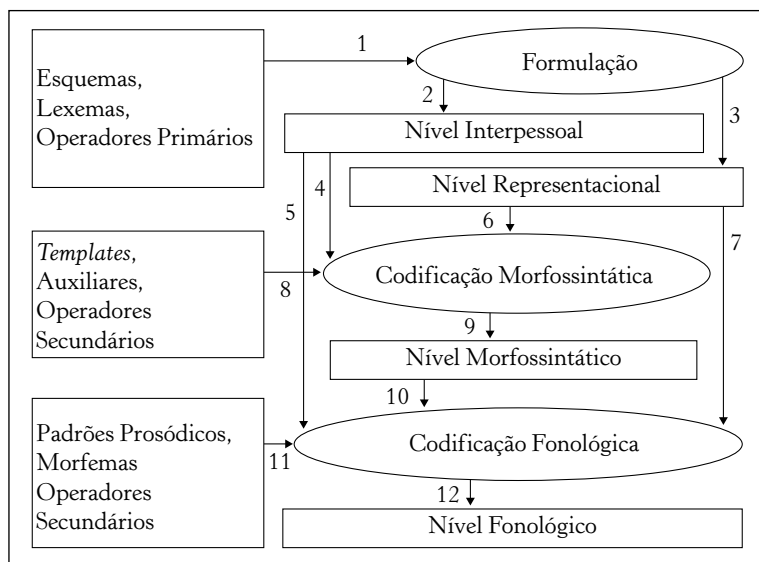
A Figura 9 representa a relação entre a organização modular da GDF e a produção, com os percursos possíveis que o usuário pode trilhar através da gramática.

(Hengeveld, 2005, p.75)

Figura 9 – Percursos através da gramática

Segundo Hengeveld (2005), as flechas horizontais sinalizam o modo como as várias operações são alimentadas pelas categorias que constituem os respectivos conjuntos de primitivos, enquanto as linhas verticais sinalizam o modo como a gramática é dinamicamente implementada durante a produção.

A noção de implementação dinâmica, postulada por Hengeveld (2005) para representar a ideia de como a gramática pode acelerar a conexão entre os diferentes níveis que a constituem, requer dois princípios relacionados à produção: o princípio da primazia da

profundidade (*depth first principle*) e o princípio da profundidade máxima (*maximal depth principle*). O princípio da primazia da profundidade, proposto originalmente por Bakker (1999; 2001) para explicar o componente da expressão, recebe uma interpretação diferente na GDF. Com efeito, o princípio da primazia da profundidade afirma que "a informação de certo nível é enviada para um nível mais baixo tão logo estiver completo o *input* informacional necessário para esse nível",[1] enquanto o princípio da profundidade máxima assegura que "somente os níveis de representação relevantes para a construção (de certo aspecto) do enunciado são usados na produção desse (aspecto do) enunciado"[2] (Hengeveld, 2005, p.73).

Um modelo de produção gramatical é tanto mais eficiente quanto mais se assemelhar à produção linguística de um indivíduo falante. De acordo com esse postulado, a informação de um nível da gramática é enviada a um nível mais baixo logo que já esteja completada no nível superior a informação *input* requerida por esse nível mais baixo.

A gramática teria um processamento consideravelmente lento, se as informações requeridas pelo Nível Interpessoal, que é o mais alto, tivessem que ser completamente especificadas em primeiro lugar e, em seguida, do mesmo modo, as do Nível Representacional, de forma que somente então a configuração morfossintática seria determinada e, em seguida, projetada sobre configuração fonológica. Hengeveld entende que não é desse modo que opera a produção linguística no indivíduo e, em consequência do postulado básico mencionado, esse modo não conduziria também a um modelo muito eficiente de gramática.

Para ilustrar esse princípio, considere-se o efeito da especificação de um valor ilocucionário no Nível Interpessoal. Nesse nível,

1 Cf. o original: *"information from a certain level is sent down to a lower level as soon as the necessary input information for that level is complete"*.

2 Cf. o original: *"only those levels of representation that are relevant for the building-up of (a certain aspect of) and utterance are used in production of that (aspect of the) utterance"*.

a seleção de um esquema imperativo para o ato discursivo produz consequências potencialmente relevantes em todos os níveis subsequentes de representação:

(i) no Nível Representacional, deverá haver a designação de um evento [+controlado] para o esquema imperativo, e o primeiro argumento deverá incluir o destinatário;
(ii) no Nível Morfossintático, algumas línguas ativam o uso de uma ordem específica para os constituintes, de auxiliares especiais de imperativo ou de marcadores morfológicos;
(iii) no Nível Fonológico, pode haver o uso de padrões prosódicos específicos para a expressão do imperativo.

Isso tudo significa que a seleção de um esquema de imperativo no Nível Interpessoal pode disparar uma grande variedade de especificações nos níveis subsequentes, tanto em termos de formulação quanto de codificação, independentemente da especificação de outros elementos no Nível Interpessoal e nos níveis mais baixos. Significa também que não é preciso completar a especificação desses outros elementos para que o esquema imperativo dispare todas as especificações mencionadas nos níveis subsequentes.

Já o princípio da profundidade máxima assegura que somente são usados os níveis de representação relevantes para a construção de um dado aspecto de um enunciado na produção desse aspecto. Esse princípio acelera a implementação da gramática por evitar a especificação vazia de níveis de representação irrelevantes à produção do enunciado em questão.

Retomando a ilustração do imperativo, nos casos em que o esquema diretivo tem que ser projetado sobre um molde (*template*) específico de oração, a aplicação dessa regra significa que, numa dada língua, pode haver uma conexão direta entre os níveis Interpessoal e Morfossintático, sem passar pelo Representacional. Similarmente, pode haver uma conexão direta entre o Nível Interpessoal e o Fonológico, sem passar pelos níveis Representacional e Morfossintático, o que pode ocorrer quando o esquema de imperativo é projetado diretamente sobre um padrão prosódico específico.

Poupam-se, desse modo, etapas supérfluas na passagem de informação no percurso descendente. De uma perspectiva ascendente, a implementação dinâmica implica que a expressão de estruturas subjacentes baseia-se potencialmente na informação advinda de todos os níveis mais altos, não apenas do nível mais alto imediatamente anterior (cf. Hengeveld, 2005).

Em vista dos dois princípios discutidos, nos percursos possíveis através da gramática, representados na Figura 9, as flechas horizontais 1, 8 e 11 se referem à consulta dos conjuntos de primitivos usados nas várias operações. Já a implementação dinâmica do modelo é vista verticalmente. Hengeveld (2005) fornece alguns exemplos de possíveis percursos através da gramática:

- 1 → 2 → 5 → 11 → 12: usa-se esse percurso quando, por exemplo, se expressa o valor ilocucionário de um ato discursivo por mecanismos prosódicos. Nesse caso, o Nível Representacional e o Morfossintático não exercem papel algum no processo de geração.
- 1 → 2 → 4 → 8 → 9 → 10: similarmente, usa-se esse percurso para projetar o valor ilocucionário de um ato discursivo sobre o molde morfossintático específico da sentença; contorna-se, nesse caso, o Nível Representacional.
- 1 → 3 → 6 → 8 → 9 → 10: emprega-se esse percurso parcial quando, por exemplo, um esquema de predicação de dois lugares tem que ser projetado sobre um molde transitivo de oração, caso em que é irrelevante o Nível Interpessoal.
- 1 → 2 → 7 → 11 → 12: usa-se esse percurso nos casos em que uma distinção semântica é expressa por meio de um padrão prosódico distinto.

Assim, depois que o tipo semântico de complemento aciona a seleção de uma determinada construção dependente, o falante precisa reacessar o Componente Contextual para verificar a disponibilidade de referentes à disposição do ouvinte, para, então, codificar a expressão no Nível Morfossintático.

Conforme prediz o princípio de primazia da profundidade, os casos de anáfora zero que são semanticamente predeterminados, como os contidos em (5-6), são motivados pelo seguinte percurso através da gramática 1→ 3 → 6 → 8 → 9 → 10, caso em que se contorna o Nível Interpessoal.

(5-6) *e eles conseguem chegar... a é óbvio uma* <u>*evolução*</u> *[****deles****] certo?* (EF-SP-405:57)

Por outro lado, os casos de anáfora zero, que são pragmaticamente determinados, como os contidos em (5-7), são ativados pelo seguinte percurso: 1→ 2 → 4 → 8 → 9 → 10; portanto, agora é o Nível Representacional que é contornado.

(5-7) *quanto à coleta se eles dependem... da colheita... de...* <u>*frutos... raízes*</u>*... que eles NÃO plantavam... que estava à disposição deles na natuREza... eles também tinham que obedecer o ciclo::... vegatativo... então existe uma época para ter uma maçã outra época para ter laranja outra época para ter banana... existem CERtas regiões onde há determinados frutos Outras regiões... com Outros frutos... então eles tinham que acompanhar este* <u>*movimento*</u> *Ø [=****dos frutos, raízes****] também:: e por isso eram nômades e não se fixavam... a lugar nenhum...* (EF-SP-405)

Observe que o referente do argumento de *movimento* – frutos, raízes – ficou estocado na memória de curto prazo do interlocutor e, quando é recuperado, no encerramento do subtópico sobre o caráter nômade do homem pré-histórico, o falante apela para essa informação compartilhada para enunciar *movimento* com um zero anafórico.

Após impor a seleção do lexema apropriado com seu respectivo esquema de predicado a (5-6), o Nível Representacional determina a forma de zero anafórico para o primeiro argumento da construção encaixada. Contudo, quanto a (5-7), a forma de zero anafórico para o primeiro argumento não é semanticamente predeterminada, mas

é motivada por razões de natureza pragmática, ou seja, pelo *status* informacional dos referentes no discurso.

Considere agora o exemplo contido em (4-26) aqui repetido como (5-8a).

(5-8) a *então a arte pré-histórica só vai poder refletir::... então a arte vai nascer:: em função dessa Necessidade... de se manter vivo... necessidade que vai se caracterizar de forma PRINcipal:: em termos de comida... isto é de caça... que é o que oferece... uma* <u>*resistência*</u> *[****da caça****] [a****o homem do paleolítico superior****] porque a:: fruta está lá então eles não precisavam se preocupar...* (EF-SP-450)

b *BEM... então vamos reconstruir a maneira de vida desse POvo para depois poder entender como surgiu a arte... e... por que surgiu um determinado estilo de arte... a gente vai pensar no* <u>*homem do paleolítico superiOR*</u>*... como um homem que ainda não conseguiu se organizar socialmente nem politicamente...* (EF-SP-450)

Um dos argumentos de *resistência* aparece no mesmo contexto sintagmático e, como tal, não pode ser expresso por ser compartilhado com o sujeito do verbo *oferecer*, representado pelo relativo retomando o nome *caça*. Já o outro argumento – *o homem do paleolítico superior* –, que já havia sido apresentado bem antes no cenário discursivo, pode ser especificado formalmente, conforme mostra (5-8b).

Por isso, defendo o princípio de que, na formulação, são os dois níveis que atuam para produzir (5-8a). Entretanto, dado que os dois níveis atuando na formulação são independentes entre si, o percurso através da gramática é exatamente o mesmo que aqueles concernentes a (5-6) e (5-7), mas em relação a (5-8) primeiramente opera o Nível Interpessoal para depois operar o Representacional para produzir ambos os tipos de zero anafórico: um referente – *homem do paleolítico superior* – é pragmaticamente motivado e o outro – *caça* – é motivado por predeterminação semântica.

5.2 A natureza aberta da predicação nominalizada

Ao analisar uma amostra de nominalizações do português falado, apresentei, no Capítulo 4, evidência empírica em apoio ao postulado de que a ausência de valência manifesta não significa que elas sejam destituídas de estrutura argumental. A Seção 5.1 mostra como argumentos não especificados, motivados semântica e pragmaticamente na formulação, são implementados como zeros anafóricos na codificação morfossintática. Nesta seção, pretendo defender a ideia de que os termos não especificados permitem alinhar as nominalizações com outros tipos de construções encaixadas não finitas em termos da natureza aberta ou fechada dos predicados e, portanto, em termos de valência potencial.

Vimos, na Seção 3.2, que segundo Dik (1997) o único traço que as nominalizações têm em comum com as construções infinitivas e participiais é a possibilidade de constituírem o núcleo de uma construção encaixada. Há, todavia, um traço adicional, igualmente relevante, compartilhado por todos os tipos de construções encaixadas não finitas e até mesmo por construções encaixadas finitas, que é a possibilidade de se realizarem como predicações fechadas ou abertas, conforme a natureza especificada ou não especificada dos argumentos disponíveis na estrutura valencial deles, o que, como vimos, não significa ausência de valência.

Dik atribui esse outro traço definidor somente a construções infinitivas e participiais, mas eu defendo que ele seja estendido às nominalizações, que são predicados valenciais como os outros dois tipos de construções encaixadas, os infinitivos e os particípios. Ao tratar de construções encaixadas, Dik (1997) não faz qualquer referência à natureza aberta ou fechada das nominalizações atuando como núcleo de predicações, que acaba sendo uma propriedade relevante para explicar a natureza valencial delas.

Vale lembrar que, conforme discutido na Seção 3.1, predicações abertas são estruturas subjacentes à espera de serem preenchidas por algum termo, de acordo com o seguinte esquema geral, mencionado em (3-12) e aqui repetido como (5-9):

(5-9) $(\Omega x_i\,(x_i: \Phi_1(x_i): \Phi_2\,(x_i): \ldots : \Phi_n\,(x_i))$

(cf. Dik, 1989, p.55)

Nesse esquema, em que Ω representa um ou mais operadores de termos e x_i simboliza o referente pretendido, cada expressão $\Phi_1(x_i)$ representa uma predicação aberta em x_i, ou seja, um "restritor" para os valores possíveis de x_i. Adicionalmente, uma predicação aberta em x_i pode representar também um esquema de predicado em que todas as posições, com exceção da ocupada por x_i mesmo, tenham sido preenchidas por estruturas de termo.

Os exemplos contidos em (5-10) e (5-11) constituem boas ilustrações de predicações fechadas, isto é, predicações com todos os argumentos especificados, tendo como núcleo um infinitivo e um particípio, respectivamente.

(5-10) *Passei sem me ver**em.*** (cf. Dik, 1989, p.55)

(5-11) *Ele chegando a São Paulo, fomos para o hotel.*

Do mesmo modo, pode-se dizer que o exemplo contido em (5-12) constitui, agora, uma boa ilustração de uma predicação fechada: o nome deverbal que atua como núcleo da construção encaixada tem todas as posições argumentais devidamente preenchidas.

(5-12) *Maria lamentou a demissão **de Pedro por João.***

Quando uma construção infinitiva ou participial expressa uma proposição em que o sujeito da construção encaixada e o da oração matriz são idênticos e, portanto, aparecem não especificados na expressão, elas são chamadas predicações abertas, conforme se observa em (5-13a-b).

(5-13) a ***Ele**$_i$ quer **ø**$_i$ trabalhar na Universidade.*
b ***Ø**$_i$ chegando a São Paulo, fo**mos**$_i$ ao hotel.*

Nas predicações abertas de (5-13a) e (5-13b), não há um constituinte especificado na representação do sujeito encaixado, já que esse constituinte, um caso comum de anáfora zero, fica não especificado sob condição de correferência com o sujeito mais alto (Dik, 1997, p.148). Embora Dik (1997) não mencione nominalizações, é verdadeiro que o mesmo processo se aplica a elas. Considere o exemplo contido em (5-14) em que a especificação do argumento A[1] de *movimento* fica automaticamente bloqueada por razões de correferência com o SN *animais* da predicação matriz:

(5-14) *consegue mostrar os animais:: em pleno movimento...* (EF-SP-405:57)

Segundo Dik (1997), construções infinitivas abertas não se restringem às que contêm um termo anafórico. Ele menciona um uso de predicação aberta nucleada por infinitivo cujo sujeito tem uma leitura genérica e não anafórica, conforme o exemplo contido em (5-15).

(5-15) *É perigoso nadar naquele lago.*

Como não há nenhum antecedente com o qual o sujeito da oração encaixada possa estar ligado, o sujeito não especificado não tem um valor anafórico; em vez disso, ele expressa que o perigo de que se fala pode aplicar-se a qualquer pessoa e, por isso, o sujeito não especificado tem uma leitura genérica e não correferencial. Uma vez mais, o mesmo se aplica a nominalizações, como mostra o exemplo (5-16):

(5-16) *É comum a demissão de trabalhadores no final do ano.*

Em suma, o exame de casos dos três tipos de construções encaixadas postulados por Dik (1997) permite deduzir que a propriedade aberta ou fechada da predicação, aplicada a construções encaixadas com infinitivos e particípios como núcleo, é perfeitamente extensível a nominalizações.

Em geral, os casos de nominalização que configuram predicações abertas têm posição de sujeito não especificada, sob a condição de correferência com o sujeito da predicação matriz ou sob a condição de interpretação genérica. Essa situação é exatamente a mesma das construções nucleadas por formas infinitivas e participiais.

Pelas razões expostas com base na evidência acima, penso ser possível reorganizar, na Figura 10, o quadro proposto por Dik (ver Figura 5), para que seja possível inserir nele uma nova modalidade de construção encaixada, as construções nominalizadas.

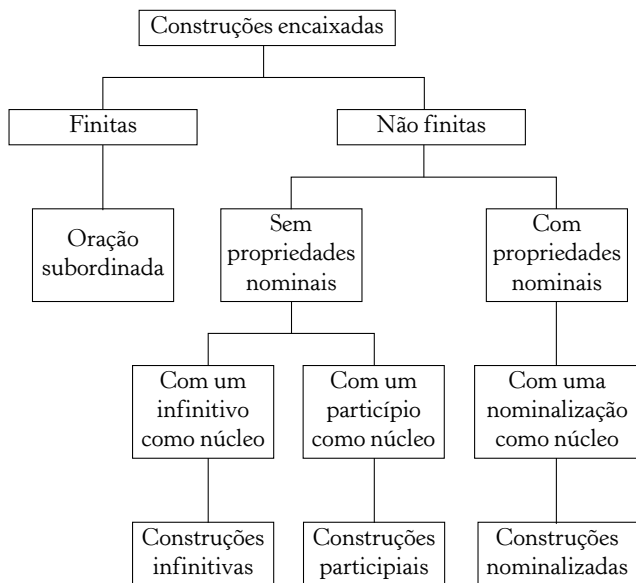



Figura 10 – Tipos de construções encaixadas revisitados

Além disso, as motivações semânticas e pragmáticas para os argumentos não manifestos podem ser explicadas pelo processo de implementação dinâmica requerido pelo modelo da GDF. Como um ponto final, é conveniente acrescentar que há uma vantagem teórica importante para esse modo de enfocar as nominalizações

e outras construções encaixadas: ele preserva a correspondência intuitiva, quase universal, entre nomes derivados e verbos de origem, o que assegura alto grau de adequação entre teoria e dados empíricos.

5.3 A articulação entre lexemas e esquemas de predicação na representação das nominalizações

Como os dados discutidos mostram a plausibilidade da hipótese de preservação de valência, é óbvio que nomes primários e derivados devem já vir selecionados do léxico com sua própria estrutura valencial. As motivações semânticas e pragmáticas acionam a forma de expressão no Nível Morfossintático. Sendo assim, é preciso considerar, em primeiro lugar, como as nominalizações entram na formulação a partir de sua organização lexical.

Quem primeiramente reconheceu que a nominalização é um tipo ambíguo de construção, a meio caminho entre duas categorias semânticas, no Nível Representacional, ou seja, entidade (x) ou predicado (f), no âmbito da GF foi Santana (2005). Ao considerar as duas hipóteses aqui discutidas, a hipótese de preservação de valência e a hipótese de redução de valência, essa autora postula que as duas posições teóricas, sustentadas respectivamente por Dik (1985, 1997) e por Mackenzie (1985, 1996), só são de fato alternativas se forem consideradas no âmbito da GF, com base na qual foram geradas. Entretanto, se forem consideradas no arcabouço da GDF, a questão se resolve de outra maneira.

Para tanto, Santana (2005) leva em conta a proposta, já discutida na Seção 3.5, avançada por Garcia Velasco e Hengeveld (2002). Vale lembrar que, ao discutirem o papel dos esquemas de predicado na GF, esses autores propõem a substituição desse conceito pelo de esquema de predicação, no interior do qual os lexemas listados no léxico, mas desprovidos de esquemas de predicado, são inseridos mediante a aplicação de regras de ligação. A ideia de introduzir es-

quemas de predicação na GF foi proposta originalmente por Hengeveld (1992 apud Garcia Velasco; Hengeveld, 2002) como uma alternativa para a Regra de Formação Termo-Predicado de Dik (1980 apud Garcia Velasco; Hengeveld, 2002).

Vale lembrar que o modo de resolver o problema que a noção de esquema de predicado levanta, com respeito aos padrões de adequação psicológica, pragmática e tipológica, é dividir esquemas de predicado em predicados e esquemas, e ter um mecanismo de ligação que os associe na gramática. Conforme já mencionado também, os autores dão preferência por usar o termo *esquema de predicação* para *esquema* e *lexema* para *predicado*, liberando o uso do termo *predicado* para se referir a lexemas usados como predicados em contextos sintáticos particulares.

Uma forte motivação para essa separação entre lexemas e esquemas de predicado é o fenômeno da especificação ou não especificação de argumentos na função de objeto, aspecto do Nível Morfossintático que tem vinculação com a nominalização. Vale lembrar também que Garcia Velasco e Portero Muñoz (2002) alegam que a não especificação de objetos é determinada por restrições de tipo semântico e de tipo discursivo.

Santana (2005) argumenta que a motivação pragmática é verdadeiramente aplicável aos casos de nominalização que permitem a recuperação anafórica dos referentes no contexto, sem, contudo, admitir mudança semântica no tipo de estado de coisas que representam. Na leitura genérica, as nominalizações acompanham o mesmo formato de verbos como *comer,* em cujo esquema não tem pertinência especificar argumentos (ver p.146). O mesmo se aplica aos casos de predicados nominais de dois lugares, em que o agente, por exemplo, é uma noção implícita que não necessita de expressão formal.

Com base nessa proposta, Santana (2005) postula que a ausência de especificação formal para os argumentos de nominalizações, como as fornecidas em (5-17a-b), pode receber o mesmo tratamento que Garcia Velasco e Hengeveld (2002) devotam à relação entre itens lexicais e esquemas de predicação:

(5-17) a *toda e qualquer manifestação que a gente for procurar vai ter que estar necessariamente ligada... a esta* ***preocupação*** *vital do homem pré-histórico de... se conservar vivo...* (EF-SP-405:50)

b *então numa vida desse tipo... a* ***preocupação*** *principal está centrada na sobrevivência...* (EF-SP-405:50)

As duas casas valenciais de *preocupação* em (5-17a) estão formalmente especificadas, situação oposta do que ocorre com a mesma nominalização em (5-17b). Entende Santana (2005) que é possível aplicar à nominalização o esquema de predicação proposto por Garcia Velasco e Hengeveld (2002), situação que produziria a representação formal de (5-18a) para (5-17a) e de (5-18b) para (5-17b).

(5-18) a preocupação [N]
$[f_1$: [Zero (x_i) [Referência (x_j)]]]

b preocupação [N]
$[f_1$: $Pred_N]$

Santana alega que, caso seja informacionalmente relevante especificar formalmente os dois argumentos de *preocupação*, a representação respectiva corresponderia a um esquema de predicado bivalente; caso não seja relevante a especificação formal de valência, a representação respectiva corresponderia a um esquema de predicação avalente.

O mesmo se aplica a um caso de nominalização com leitura genérica em que os participantes não são formalmente especificados como se vê no exemplo grifado de (5-19)

(5-19) *nessa época ainda não existe preocupação com composição...* (EF-SP-405)

Na perspectiva de Santana, esse caso é similar a verbos intransitivos como *comer,* em que o objeto só aparece quando especificado.

Assim, na visão de Santana (2005, p.151), tanto o tipo semanticamente motivado de zero do nome deverbal quanto o tipo pragmaticamente motivado receberiam um esquema de predicado de zero lugar, conforme as próprias palavras da autora:

> As evidências já discutidas anteriormente sugerem que a ligação entre lexemas e esquemas de predicação não pode ser desenvolvida numa base biunívoca; pelo contrário, os lexemas podem assumir diferentes esquemas, se eles tiverem mais de uma leitura, como é o caso de *open* ("abrir"), como predicado intransitivo e como predicado transitivo. Estendendo mais os fatos a fenômenos similares, neste trabalho, postula-se que essa abordagem seja aplicada, nos mesmos termos, aos predicados nominais derivados, com a evidente vantagem de dar caducidade à necessidade de escolha entre a *hipótese de preservação de valência* e a *hipótese de redução de valência*.

A solução proposta por Santana (2005) é similar, mas não idêntica à que postulo aqui, já que ela defende que os predicados genéricos e os não referenciais teriam um esquema de predicação avalente. Os predicados referenciais, cujos argumentos estivessem não especificados formalmente, receberiam um esquema de predicado de zero, um ou dois lugares, conforme aparecem na expressão. Minha interpretação diverge no sentido de que a genericidade e a referencialidade, de que resulta apenas não especificação formal de argumentos no Nível Morfossintático por razões de economia, são traços do Nível Interpessoal uma vez que a apresentação das entidades no discurso é responsabilidade dos participantes da interação; quanto aos predicados referenciais, a motivação pode ser tanto semântica quanto pragmática, conforme visto na Seção 4.4.

Embora não concorde com a interpretação de Santana, concordo com a ideia de que itens lexicais e esquemas de predicados devem ser entidades separadas. No máximo, *composição* em (5-19), aqui repetido como (5-20a), deve receber um esquema de predicado com duas posições abertas como aparece em (5-20b), já para o nome

preocupação em (5-20a), o esquema de predicado deve ter uma posição estrutural fechada e uma aberta como em (5-20c).[3]

(5-20) a *nessa época ainda não existe preocupação com* <u>*composição*</u>... (EF-SP-405)

b $(e_i$: $[(f_i$: composição $(f_i$:$)$ $(x_i)_A$ $(x_j)_U)]$ $(e_i))$

c $(e_i$: $[(f_i$: preocupação $(f_i$:$)$ $(x_i)_U$ $(x_j$: composição $(x_j)_A]$ $(e_i))$

Nesse caso, o esquema de predicado é de fato aplicado *online* conforme as necessidades da situação discursiva, mas a valência deve vir marcada na representação do item no próprio léxico. O Nível Morfossintático é responsável pela indicação do zero anafórico como forma de expressão para o argumento experienciador de (5-20), conforme a sugestão de interpretação fornecida em (5-21), em que $Ø_i$. mostra que o núcleo do SN é um zero coindexado a algum outro constituinte já mencionado no discurso.

(5-21) $[[[\text{preocupar}]_V\text{ção}]_N\ [Ø_{i]SN}[\text{com}_P\ \text{composição}]_{SP}]_{SN}$

O problema com a solução fornecida por Santana (2005) é que ela é contraditória com as evidências empíricas, já mostradas na análise dos dados. Essas evidências mostram que a especificação formal de valência na estrutura do nome nuclear depende de um conjunto de fatores semânticos e pragmáticos, relacionados principalmente ao compartilhamento de informação de curto prazo pelos participantes do processo de interação verbal.

Mais grave ainda, essa alternativa acaba sendo contraditória com a própria defesa da hipótese de preservação de valência sus-

3 Os elementos subscritos A e U representam nos exemplos a seguir as funções semânticas de Actor (Ator), Undergoer (Paciente), que se referem, respectivamente, a argumentos de um estado de coisas dinâmicos com uma participação mais ativa e mais passiva. Outro elemento subscrito é L para Locative (Locativo) que tem a função de indicar o lugar de ocorrência do estado de coisas, seja ele dinâmico, seja não dinâmico (cf. Hengeveld; Mackenzie, 2008, p.194-204).

tentada por Santana (2005). A configuração formal adotada pela autora não acompanha o princípio teórico de que a representação subjacente das nominalizações é dotada de uma estrutura argumental cuja especificação no dado superficial depende de regras derivadas do uso. Como visto, essas regras podem ser de natureza pragmática em função de determinações do Nível Interpessoal, ou podem ser de natureza semântica em função de determinações do Nível Representacional. Vale destacar que, na literatura funcionalista, essas motivações estão ligadas aos princípios de economia e de iconicidade (Givón, 1980, 1990; Haiman, 1983, retraduzidos por Cristofaro, 2003).

Na análise apresentada neste trabalho, sustento portanto que, se a estrutura argumental permanece preservada, tanto o predicado verbal quanto o predicado nominal devem receber um esquema de predicado com a mesma estrutura argumental: a inserção do lexema respectivo – nome ou verbo – depende do léxico.

Concordo plenamente com Garcia Velasco e Hengeveld (2002) que esquemas de predicação devem ser investidos na gramática e, portanto, ser separados do léxico, conforme visto na Seção 3.4. Essa separação, que torna obsoletas as regras de derivação lexical previstas por Dik (1989), torna o léxico mais econômico. Já no que se refere à especificação de argumentos, o esquema de predicado pode representar, no máximo, uma predicação aberta com uma, duas ou três posições, mas deve conter tantas posições estruturais quantos forem os argumentos do item lexical selecionado.

Para fornecer uma representação formal simplificada dessa interpretação nos níveis Interpessoal e Representacional, considerem-se os exemplos de (5-22a) e (5-23a) e os esquemas de predicação respectivos de (5-22b) e (5-23b), já representando a interpretação de que as nominalizações são propriedades configuracionais (f) no Nível Representacional e Subatos de Atribuição (T) encaixados em Subatos de Referência (R) no Nível Interpessoal. A valência quantitativa é variável de acordo com a natureza do item que for escolhido no léxico, o que permite a expressão de (5-22a) e (5-23a) de modo similar à dos predicados verbais correspondente.

(5-22) a *A manifestação dos caras-pintadas foi uma das causas da renúncia de Collor.*

b $(R_i\ [T \quad R_j \quad]\ (R_i))$

$(e_i\text{: } [(f_i\text{: } \blacklozenge\ (f_i\text{:})\ \ (x_1)_U)] \quad (e_i))$

(5-23) a *A invasão de Cartago pelos romanos mudou o comércio do Mediterrâneo.*

b $(R_i\ [T \quad R_j \quad R_k\ ] \quad (R_i))$

$(e_i\text{: } [(f_i\text{: } \blacklozenge\ (f_i\text{:})\ \ (x_i)_A \quad x_j)_U)\ ] \quad (e_i))$

Esses esquemas expressam que os lexemas *manifestação* e *invasão* designam uma relação (aqui representada pela variável f) entre uma entidade no caso de (5-22a) e duas entidades no caso de (5-23a) (representadas pelas variáveis x). A presença dessas duas variáveis na representação do lexema do nome, segundo Velasco e Hengeveld (2002, p.114), guia o processo de ligação na direção de um esquema de predicação de um lugar para (5-22a), e de dois lugares, para (5-23a).

Observe ainda a natureza da correspondência entre o Nível Interpessoal e o Nível Representacional. Nos dois casos, trata-se, no Nível Interpessoal, de um Subato de Referência, que se faz mediante o uso de nominalizações, que tem sob seu escopo, todavia, um Subato de Atribuição, o que representa o caráter complexo da nominalização. Já no Nível Representacional, trata-se de um estado de coisas (e) que tem sob seu escopo uma propriedade (f) que, por sua vez, estabelece uma relação semântica entre uma ou duas entidades (x).

5.4 Uma proposta de representação formal para a nominalização

Outra questão que se coloca naturalmente é como considerar, no arcabouço da GDF, o caráter híbrido das nominalizações, a meio caminho entre a referência a uma entidade de primeira ordem,

como os nomes comuns concretos, e a referência a um estado de coisas, como entidades de ordem superior. A título de exemplificação, considere-se o predicado de dois lugares *invasão* em (5-23a) acima.

Também para efeito de simplificação, considere-se a representação subjacente da nominalização, encaixada na posição de sujeito da predicação matriz. Para dar conta do caráter categorial híbrido da nominalização é necessário considerar que ela representa, no Nível Interpessoal, um Subato de Atribuição, próprio de uma predicação, dentro de um Subato de Referência, próprio de uma entidade, tal como se representa em (5-24b), e, no Nível Representacional, representa a atribuição de uma propriedade (f) a duas entidades referenciais (*Cartago* e *romanos*), que constitui a predicação encaixada numa proposição em (5-24c).

(5-24) a *A invasão de Cartago pelos romanos*
b $(R_i\ [T_i \quad (R_j) \quad (R_k)\ ] \quad (R_i))$
c $(e_i\ [(f_i\text{: invasão})_N\ (f_i)\ \ (x_i\text{: romanos}(x_i)_A\ \ (x_j\text{: Cartago}\ (x_j)_U]\ \ (e_i))$

Considerando o estatuto comunicativo das nominalizações na GDF, se consistem num Subato de Atribuição ou num Subato de Referência, Mackenzie (2004a) aponta para o fato de que operadores de identificabilidade e especificidade se tornam mais acessíveis conforme aumenta o estatuto nominal da predicação encaixada. Operadores de indefinitude são também passíveis de atuar sobre o Subato de Referência, conforme se vê em w*inning this battle is a winning of all battles* e *there won't be any winning of hearts and minds* (Mackenzie, 2004a).

Assim, para fazer justiça ao caráter ambíguo das nominalizações a melhor análise é considerar, como faz Mackenzie (2004a), que o lado mais nominal da predicação nominalizada obriga a representá-la como um Subato de Referência que abriga um Subato de Atribuição. Assim, nos casos em que a nominalização não admite operadores de definitude, a leitura é de uma predicação com sua própria estrutura argumental (5-25a) e, nos casos em que a nominalização admite esse tipo de operador, a leitura é de um Subato de

Referência dentro de um Subato de Atribuição sem estrutura argumental, de acordo com a hipótese de redução de valência (5-25b):

(5-25) a *I* *enjoyed* *my horse* *winning* *the race.*
R T R T R
(e_1: [(f_1: win (f_1) (my horse)$_A$ (race) $_U$] (e_1))

b *Everyone* *witnessed* *my horse's* *winning* *of the race*
R T (d/iR: [R T R] (R)
(e_1: [(f_1: winning (f_i)) (my horse)$_{Ass}$ (race) $_{Ass}$] (e_1))

(cf. Mackenzie, 2004a)[4]

Uma diferença, contudo, persiste na análise aqui apresentada e na de Mackenzie. Como se vê em (5-25) acima, persiste em sua análise a necessidade de considerar que, quando as nominalizações passam a funcionar como um Subato Referencial no Nível Interpessoal, representado por um estado de coisas no Nível Representacional, elas perdem necessariamente a estrutura argumental que só se mantém nos casos, válidos para o inglês, em que se preserva o estatuto lexical de predicado verbal.

O modo como os argumentos dessa predicação encaixada são expressos vai depender, como se viu anteriormente, de processos de indeterminação semântica, motivados pelo grau de conexidade entre a oração da predicação encaixada e a oração da predicação matriz, ou por processos pragmáticos, motivados pelo grau de acessibilidade de informação no discurso corrente. Esses dois processos são responsáveis pelo acionamento de expressão dos argumentos por anáfora zero, mas são responsáveis também pela indeterminação de entidades argumentais ou pela ocorrência de nominalizações não referenciais. Todos esses processos de expressão argumental

4 A recusa de Mackenzie (1996), que permanece nesse trabalho de 2004a, em ver argumentos na nominalização produtiva em inglês o faz substituir as funções semânticas de Ator e de Paciente, que ele admite no predicado verbal, para a função semântica *Associativo* na nominalização.

são repercussões diretas dessas motivações funcionais no Nível Morfossintático.

Assim, para fazer justiça às motivações funcionais e ao caráter semântica e pragmaticamente híbrido das nominalizações, a expressão morfossintática deve respeitar as duas camadas de atuação nos Níveis Interpessoal e Representacional e receber também duas camadas de representação, que poderiam ser denominadas "sintaxe interna" e "sintaxe externa" em atenção à proposta de Haspelmath (1995) para lexemas derivados por flexão. O autor se refere a processos de derivação por flexão transposicional, muito produtivos num grande número de línguas, que são responsáveis pelo caráter morfossintaticamente híbrido das formas resultantes, conforme a representação que faz da forma *singende* do alemão, transcrita em (5-26).

(5-26)	der	im Wald	laut	sing$_V$-ende$_{Adj}$	Warderer
	the	in the forest	loud	sing-PTPC	hiker

'the hiker (who is) singing loud in the forest'

(Haspelmath, 1995, p.44)

É útil aplicar o mesmo padrão a outras formas híbridas como *singende*, mas produzidas por processos comuns de derivação por sufixação, do que resulta a representação morfossintática para a expressão nominal *a invasão de Cartago pelos romanos* contida em (5-27).

(5-27) [[det sing [invadir]$_v$ são]$_N$ [de Cartago]$_{SP}$] [por det pl romanos] $_{SP}$] $_{SN}$

O esquema representado em (5-28) abaixo representa o SN de acordo com os três níveis de organização previstos pela GDF:

(5-28)	$(R_i$ $[(T_i)$	(R_j)	(R_k)] $(R_i))$
	$(e_i$ $[(f_1)$	$(x_i)_U$	$(x_j)_A$] $(e_i))$

[[det sg[invadir]$_V$ são]$_N$ [de Cartago]$_{SP}$] [por det pl prep romanos] $_{SP}$] $_{SN}$

Uma das razões para nominalizar é a possibilidade de as nominalizações retomarem, por substituição lexical, outras predicações já mencionadas no discurso precedente, como se observa, por exemplo, em (5-29) e (5-30).

(5-29) *os animais iam hibernar outros... i**migra**vam para lugares mais quentes eles também precisavam acompanhar.... o a **migra**ção da caça senão eles iam ficar sem comer...* (EF-SP-405)

(5-30) ***cria**r uma pessoa... ou **cria**r uma imagem é mais ou menos a mesma coisa... no sentido de que nós estamos **cria**ndo uma coisa nova... do nada... eu não tinha nada aqui passo a ter a imagem da minha mão... e esta ideia de **cria**ção é que ainda ( ) é representação...* (EF-SP-405)

A mesma relação em dois diferentes níveis morfossintáticos pode representar a forma dos argumentos na forma de adjetivo, como se pode ver em (5-31) e (5-32).

(5-31) *então nós vamos começar pela Pré-História... hoje exatamente pelo período... do paleolítico... a **arte**... no período paleolítico [...] as:: manifestações **artísticas** começaram a aparecer no paleolítico superior* (EF-SP-405)

(5-32) *no final das contas toda a evolução **humana**... não deixa de ser exatamente a evolução do domínio que o **homem** tem sobre a natureza...* (EF-SP-405)

A manifestação de A^1 sob a forma de adjetivo mostra que a representação morfossintática deve acompanhar a recuperação lexical de um nome mencionado no discurso precedente como em (5-31); todavia, o adjetivo pode ser retomado como nome no discurso subsequente, como (5-32). Outro caso interessante foi registrado por Keizer (2004b), que reproduzo aqui em (5-33):

(5-33) *In Paris Francine Stock examines the thinking behind the last minute French initiative what they hope do gain from it.* <ICE--GB:S2B-010 #7:1:1>

(Keizer, 2004b, p.12)

O adjetivo *French* atua como um verdadeiro argumento tanto semântica quanto pragmaticamente e, como tal, é interessante verificar que pode ser retomado por um pronome, que é, por definição, uma classe que substitui o nome.

Essas possibilidades de intercâmbio entre as categorias lexicais nas relações de substituição lexical consistem num forte argumento para considerar os adjetivos como expressão argumental similar a nomes, e não como meros modificadores, relação mais apropriada para nomes de primeira ordem, como *humano* em *utensílio humano*.

As evidências de natureza textual-discursiva sugerem que a ligação entre lexemas e esquemas de predicação deve ser desenvolvida tanto com base na sintaxe interna quanto com base na sintaxe externa; na sintaxe interna, um nome deverbal ocupa apenas provisoriamente a posição estrutural de um verbo, assim como um adjetivo, a de um nome, enquanto na sintaxe externa trata-se de nome, na função de núcleo de um sintagma nominal e de um adjetivo, na função de modificador. O ajuste a que se refere Dik (1985; 1997) dos nomes derivados de ordem superior ao nome prototípico de primeira ordem se aplica, portanto, somente à sintaxe externa.

6
O LUGAR DA NOMINALIZAÇÃO NO *CONTINUUM* CATEGORIAL

Introdução

Os dados analisados mostraram que as nominalizações no português preservam a estrutura argumental do verbo *input* correspondente e que a expressão formal na estrutura do nome nuclear depende de um conjunto de fatores semânticos e pragmáticos, relacionados principalmente ao compartilhamento de informação de curto prazo pelos participantes do processo de interação verbal. Como já mencionado, na literatura funcionalista essas motivações estão ligadas aos princípios de economia e de iconicidade (Givón, 1980, 1990; Haiman 1983).

O próximo passo no percurso aqui adotado para demonstrar a tese de que a continuidade categorial das classes de palavras se aplica aos nomes e aos verbos do português é postular motivações funcionais para a explicação da ausência de especificação de argumentos como evidência da natureza complexa da nominalização; em seguida, demonstrar a tese do caráter prototípico de nomes e verbos em relação à nominalização.

Essa discussão se organiza da seguinte maneira: na Seção 6.1, desenvolvo o princípio teórico, com base em evidência empírica, de que a nominalização constitui uma entidade complexa, meio

híbrida, a meio caminho entre o nome e o verbo prototípicos; na Seção 6.2, faço uma análise de nomes e verbos do português com base na Teoria dos Protótipos; na Seção 6.3, mostro que há uma relação de simetria funcional entre a classe das palavras referenciadoras e a classe das palavras predicadoras; finalmente, na Seção 6.4, desenvolvo a ideia especulativa de que o português se enquadra tipologicamente na classe das línguas diferenciadas pelo modo como organiza a realidade percebida em classes de palavras.

6.1 A natureza complexa da nominalização

Conforme discutido no Capítulo 1, o critério semântico de estabilidade temporal, invocado por Givón (1979) para determinar a prototipicidade dos nomes, não é suficiente para atribuir uma classe lexical a uma dada forma: "A prototipicidade de categorias linguísticas depende não apenas de propriedades semânticas independentemente verificáveis, mas também – e talvez mais crucialmente – da função linguística no discurso" (Hopper; Thompson, 1984, p.708).[1] Nem sempre o uso de um nome é capaz de construir ou de identificar um referente (Dik, 1989, p.114), como ocorre, por exemplo, com o termo *definição* na sentença (6-1).

(6-1) *ela quer saber as matérias que ela vai ter... o curso::... o segundo ciclo que ela pretende fazer... sabe? bom já está numa:: idade de* <u>*definição*</u> *quanto ao segundo ciclo porque elas já estão na oitava série as mais velhas não é?* (D2-SP-360)

Segundo Hopper e Thompson (1984), o fato semântico de que um termo denota uma entidade concreta, visível (que já não é o caso de uma entidade de segunda ordem, como *definição*) não é crucial para

1 Cf. o original: *"prototypicality in linguistic categories depends not only on independently verifiable semantic properties, but also – and perhaps more crucially – on linguistic function in the discourse".*

determinar se esse termo é um membro prototípico da classe. Em vez disso, é mais importante e decisivo que o nome exerça algum papel no discurso em que ele figura, seja construindo seja identificando um referente na interação, como o caso de *definição* em (6-2), núcleo de um SN com um conjunto completo de possíveis constituintes.[2]

(6-2) *Você entendeu aquela segunda definição de sintagma que o professor deu na aula de ontem?*

Do mesmo modo, os traços semânticos do verbo (visibilidade, movimento e efetividade) não são suficientes para determinar sua prototipicidade. Para qualificar-se como membro prototípico da classe, uma forma verbal deve referir-se à ocorrência de um evento do discurso. Assim, uma forma não finita, como *acertar* em (6-3), não é um membro prototípico como o de (6-4), em que a forma finita representa o evento de acertar como efetivamente concluído no passado. Também para verbos, é o papel discursivo da forma empregada que consiste, para Hopper e Thompson (1984), no principal fator que determina se ele é um membro central ou periférico de sua categoria.

(6-3) *Acertar traves exige habilidade.*

(6-4) Neymar acertou a bola na trave para provar que é habilidoso.

Essas relações discursivas, que são representadas pelo Nível Interpessoal, é que determinam o grau de complexidade categorial e, portanto, o de versatilidade sintática das nominalizações, não necessariamente a redução de argumentos.

2 Considerando que entidades são coisas que podem ser construções mentais, é possível distinguir duas funções principais para o uso de termos: ao empregar um termo numa referência construtora, o falante pretende ajudar o ouvinte a construir um referente, o que implica introduzir a entidade no modelo mental do ouvinte; já ao empregar um termo numa referência identificadora, o falante pretende ajudar o ouvinte a identificar um referente que já se acha disponível (cf. Dik, 1989, p. 114).

A nominalização tem como modelo prototípico o nome comum não derivado e, como tal, usa a expressão de possuidor, típica dos nomes comuns, para a expressão argumental. Entretanto, enquanto a nominalização fizer referência não a uma entidade de primeira ordem, mas a uma entidade de ordem superior, a correspondência entre os argumentos do nome e os do verbo *input* deve estar representada na representação subjacente de ambas as classes de palavras.

Um bom argumento para sustentar uma diferença entre nomes de primeira e nomes de segunda ordem são os casos de dupla manifestação de possuidor, principalmente na estrutura do tipo genitivo-possessivo. Embora a nominalização se ajuste ao modelo dos nomes prototípicos (6-5a), há uma diferença absoluta entre eles, já que a manifestação de dois possuidores só é licenciada para as nominalizações pelo uso do genitivo em inglês (6-5b) e por um pronome possessivo em português (6-5c).

(6-5) a ** o livro de José de Maria / * John's book of Mary*
b *John's decision of attacking.*
c *Minha decisão de ataque.*

Muito raramente nomes deverbais se referem a entidades de primeira ordem, especialmente quando representarem nominais de ação, como em (6-6a). A única situação em que é possível esse tipo de referência é aquela em que os deverbais indicam estados de coisas resultantes de ação, como em (6-6b).

(6-6) a *A <u>construção</u> da casa demorou dois anos.*
b *Aquela <u>construção</u> do alto da colina é muito sólida.*

O nome destacado em (6-6b) se deriva do verbo *construir*, mas não é dotado de estrutura argumental, por não se referir a um estado de coisas, mas ao resultado de um estado de coisas, que constitui uma entidade de primeira ordem; nesse caso, a função do SP *do alto da colina* é localizar a entidade referida. Já o mesmo nome deverbal em (6-6a) representa uma predicação encaixada na posição de sujeito e a função do SP *da casa* é indicar o argumento paciente de *construção*.

Em vista dessa diferença, é possível aplicar os rótulos categoriais de (6-7) para as duas expressões possíveis do nome em comparação ao verbo:

(6-7)	construir	>	construção$_1$	>	construção$_2$
	[+ V – N]		[+V + N]		[-V + N]

Na escala de desverbalização (Malchukov, 2004), a nominalização só perde a condição de referência a um estado de coisas quando se transforma realmente num nome de primeira ordem. As categorias mais externas e respectivas camadas são mais prontamente afetadas pelas operações transcategoriais do que as mais internas. Desse modo, a valência é, na visão do autor, a camada semanticamente mais próxima do lexema verbal e, portanto, mais preservada do que as demais categorias. Considerando os processos complementares de desverbalização e de nominalização, Malchukov propõe o modelo em (6-8), que ele chama Modelo Escalar Generalizado:

O que ocorre com o nome construção$_2$ no esquema em (6-7) é que, além de descategorizar-se a partir de um lexema verbal, ele continua o processo, recategorizando-se como termo referente a uma entidade de primeira ordem. Nesse caso, perde valência e recebe todas as marcas possíveis de um membro prototípico da classe dos subtantivos. Já o que ocorre com construção$_1$ é que apenas sofre

3 As abreviaturas têm o seguinte significado: no polo nominal, N representa nome; CL, classificador; NUM, número; POS, possuidor; DET, determinante e CAS, caso. Já no polo verbal, V representa verbo; VAL, valência; ASP, aspecto; TEM, tempo; MOD, modo; CONC, concordância; FI, força ilocucionária.

processo de descategorização, o que implica preservação da referência a uma entidade de ordem superior e, consequentemente, da estrutura argumental do predicado verbal correspondente.

Malchukov não se refere a redução de valência nesse processo de descategorização, mas a efeitos de regras de bloqueio; por exemplo, no caso da forma gerundivo-genitiva do inglês, como em *John's driving the car*, o processo de derivação lexical atinge o argumento agente, ficando o paciente com a mesma marcação que mantém no verbo, isto é, como se não fosse afetado pelo efeito das regras de bloqueio.

Na realidade, as regras de bloqueio não reduzem o número de relações gramaticais da categoria verbal, mas o preservam no nome deverbal, alterando apenas o mecanismo formal de marcação dos argumentos. Nesse caso, o uso de preposições como *de* e *por* constitui um mecanismo formal das nominalizações para visibilizar as mesmas relações gramaticais de sujeito e de objeto, que, em português, pelo menos, são geralmente marcadas por outros mecanismos, tais como ordem de palavras, posição pré e pós-verbal respectivamente, e concordância verbal.

Essa interpretação encontra respaldo na proposta de classificação que Keizer (2004a) faz das preposições do inglês no arcabouço teórico da GDF. Defende a ideia de que elas podem atuar tanto como elementos gramaticais (operadores) quanto como elementos lexicais (predicados). Na condição de categoria lexical, as preposições atuam como predicados de um lugar (transitivas) e como avalentes (intransitivas). Já o uso gramatical está confinado às preposições *of* e *by* quando usadas para introduzir termos com função semântica de Agente, posicionador, Força, Processado, Zero ou Meta, isto é, as funções semânticas que podem ser atribuídas ao primeiro e ao segundo argumento de predicados deverbais ou deadjetivais.

A autora restringe a discussão ao uso dessas preposições introduzindo SNs cujos núcleos consistem em predicados nominais deverbais e deadjetivais. Indiretamente, Keizer (2004a) concorda com Dik (1989, 1997) que tanto a valência do predicado *input* quanto a função semântica dos argumentos são preservadas no processo de

derivação lexical. Na medida em que, como elementos gramaticais, as preposições visibilizam funções semânticas, o enfoque deixa claramente implícita a ideia de que núcleos nominais deverbais/deadjetivais são dotados de estrutura argumental. Nas próprias palavras da autora:

> Em outros termos, o uso gramatical de preposições é restrito aos casos em que a preposição expressa uma função semântica que, na construção verbal ou adjetival correspondente, não é expressa por meio de uma preposição mas por algum (outro) meio gramatical como ordem de palavra, concordância ou caso. (Keizer, 2004a, p.27)[4]

Na esteira da posição de Keizer (2004a), considero que, em (6-6b) acima, a preposição *de* estabelece uma relação de modificador com o núcleo nominal, devendo ser analisada como preposição lexical e a relação que estabelece especifica a localização da construção, enquanto em (6-6a) estabelece uma relação argumental, consistindo, portanto, numa preposição gramatical. A mesma distribuição se sustenta nas relações estabelecidas pelas preposições *por* e *em*, nos exemplos de (6-9a-b) e (6-10a-b), em que se percebe a correspondência entre preposição gramatical para relações argumentais e preposição lexical para relações de modificação.

(6-9) a *a destruição (da cidade)* ***[por epidemia]***
b *a destruição (da cidade) (por tropas inimigas)* ***[no ano passado]***

(6-10) a *o livro* ***[com capa vermelha]***
b *a preocupação (de Maria) (com José)*

---

4 Cf. o original: *"In other words, the grammatical use of prepositions is restricted to those cases where the preposition expresses a semantic function which in the corresponding verbal or adjectival construction is not expressed by means of a preposition but by some (other) grammatical means, such as word order, agreement or case".*

Constituem relações de modificação os SPs negritados e entre colchetes nos exemplos de (6-9) e de (6-10a). Ao incidirem sobre um predicado nominal, indicam respectivamente causa (6-9a), tempo (6-9b) e especificação de referência (6-10a). Já a mesma interpretação não se aplica aos SPs entre parênteses, que constituem de fato argumentos requeridos pela estrutura argumental do predicado.

Essa interpretação não afeta diretamente a interessante correlação entre desverbalização progressiva e redução de valência proposta por Mackenzie (1966). A correlação, nos termos aqui propostos, se estabelece nos moldes propostos por Dik (1989, 1997) para o ajuste formal dos nomes derivados, isto é, nos estágios mais avançados de nominalização, a marcação das relações gramaticais tipicamente argumentais (ordem de palavras, concordância de sujeito e de objeto e caso) passa a ser visibilizada por outros mecanismos, como uso de preposições, de genitivo, de pronomes possessivos e de adjetivos, que são as categorias formais que normalmente acompanham o nome prototípico, como em (3-23a-d), aqui reescrito como (6-11a-d).

(6-11) a *That **my horse** (A^1 sentencial) won **the race** (A^2 sentencial) came as a great surprise*. (oração finita)
b ***My horse** (A^1 sentencial) winning **the race** (A^2 sentencial) came as a great surprise*. (gerúndio)
c ***My horse's** (A^1 genitivo) winning **the race** (A^2 sentencial) came as a great surprise*. (gerúndio-genitivo)
d ***My horse's** (A^1 genitivo) winning **of the race** (A^2 sintagma de possuidor) came as a great surprise*. (nominalização produtiva)

Os tipos de nominalização contidos em (6-11a-d) se correlacionam às principais classes tipológicas de línguas propostas por Koptjevskaja-Tamm (1993). A tipologia sintática das nominalizações proposta por essa autora, e anteriormente por Comrie e Thompson (1985), distingue quatro tipos principais de categorias tipológicas de línguas de acordo com o modo como se manifestam

as relações gramaticais. As relações gramaticais são representadas por S (sujeito de intransitivo), A (sujeito de transitivo) e O (objeto de transitivo), como originalmente proposto por Dixon (1976 apud Koptjevskaja-Tamm, 2003). Assim (6-11a) e (6-11b) representam o tipo sentencial (SENT), em que os argumentos S, A e O retêm a marcação sentencial própria. (6-11c) representa o tipo possessivo--acusativo (POS-ACC), em que A e S recebem marca de possessivo e O mantém a expressão sentencial.

Na classe das línguas ergativo-possessivas (ERG-POS), O e S recebem forma de genitivo e A recebe marcação diferente. O português, juntamente com o georgiano, o russo e o abkhaz, se enquadra num subtipo dessa classe tipológica de línguas, denominadas oblíquo-possessivas (OBL-POS), em que O e S se tornam genitivos, enquanto A recebe marca de oblíquo.[5] O quarto tipo principal é o tipo nominal (NOM), em que todos os argumentos são codificados como possuidores, que está representado acima pela sentença (6-11d) do inglês, como um subtipo de possessivo duplo (DUB-POS).

Koptjevskaja-Tamm (1993) faz menção a classes menores de língua com padrões de predicação em que parece haver redução de valência. Geralmente, o argumento aparentemente omitido é incorporado à nominalização complexa (tipo INC) ou é expresso dentro de uma oração relativa referindo-se ao nominal de ação (tipo REL).[6] É difícil falar em redução de valência em casos como esses, já que a incorporação e a oração relativa parecem constituir apenas estratégias para alocar um argumento, O e A respectivamente, na sintaxe externa da nominalização; a opção de alocar constituinte valencial na sintaxe externa é comum também ao português, como ocorre com (6-12):

(6-12) *é inCRÍvel o que aparece lá os <u>cortes</u> [que eles dão nas cenas]* (DID-DP-234)

---

5 O exemplo equivalente em inglês do tipo ERG-POS é *The winning of the race by my horse was no surprise*, adaptado de Mackenzie (1996).

6 INC representa o tipo incorporativo e REL, o tipo relativo.

Dentro dos tipos POSS-ACC e do ERG-POSS, é possível distinguir construções com nomes de ação que envolvem nominalizações com marcadores manifestos de voz. Nas construções derivadas, os nomes de ação estão correlacionados a verbos destransitivizados por meio de passivização e antipassivização. Assim, o único efeito da nominalização consiste em assimilar o "sujeito" de tais verbos intransitivos a nominais possessivos em SNs não derivados.

Com base nos casos que envolvem marcadores expressos de antipassiva e de passiva, Mackenzie (1985) sugere que a redução de valência, que recebe expressão morfológica regular, sempre subjaz ao processo de nominalização (Koptjevskaja-Tamm, 1993, p.62). Conforme se pode observar, todavia, como a relação de correspondência se dá entre a nominalização e um verbo intransitivo correspondente, não o verbo transitivo, é difícil falar em redução de valência nesses casos também. É justamente por isso que construções desse tipo estão fora da classificação geral da autora.

6.2 Aplicação da Teoria dos Protótipos a nomes e verbos do português

A seção anterior voltou-se para a discussão da nominalização em relação aos membros mais prototípicos das classes do nome e do verbo. Pouca atenção até aqui foi devotada a esses representantes mais centrais da categoria.

Parece que esse tópico toca de modo crucial a questão da definição das categorias linguísticas pelas gramáticas destinadas ao ensino, onde reina mais confusão que ordem, por conta do fato de que as gramáticas encaram as classes de uma perspectiva prototípica, mas sem a exata percepção do alcance dessa noção.

É comum as gramáticas fornecerem definições baseadas nos membros mais característicos da categoria para, em seguida, inserirem exemplos de nomes abstratos que designam entidades de segunda ordem. Esse procedimento é de pronto auxiliado por subcategorizações, principalmente as de abstratos e concretos, próprios

e comuns, habitualmente aplicadas aos nomes. No entanto, parece muito mais viável tratar de membros mais prototípicos, como os que designam entidades de primeira ordem, e membros menos prototípicos, como os que designam entidades de ordem superior.

Esta seção destina-se particularmente a mostrar como seria possível detectar exemplares prototípicos e não prototípicos de nomes e verbos, as duas categorias que acabam definindo o caráter complexo da nominalização.

Conforme se viu no Capítulo 1, segundo Du Bois (1980) e Givón (1979) codificar participantes no discurso depende muito pouco ou nada da noção lógico-semântica de referencialidade, envolvendo "existência em algum mundo", o que serviria, em consequência, para estabelecer as condições de verdade de uma proposição. Entende Du Bois (1985) que um SN é referencial quando usado para falar de um objeto, com identidade contínua no tempo. Entende Givón (1979) que a codificação dos participantes depende muito mais significativamente da intenção comunicativa de quem enuncia o discurso e, especialmente, do fato de um argumento individual ter suficiente importância no discurso subsequente, isto é, tem importância se o falante deseja fornecer uma identidade específica ou uma identidade genérica ao referente.

Du Bois (1985) e Givón (1979) usam o termo "referencial" para argumentos que apresentam identidade contínua e são importantes ao discurso subsequente e "não referencial" para os que não apresentam. Hopper e Thompson (1984) concordam com os autores mencionados, mas preferem as categorias "manipulável/não manipulável", para evitar as conotações lógico-semânticas dessa oposição.

Assim, o tratamento das classes como detentores de membros mais ou menos prototípicos, principalmente o dirigido ao enfoque de nomes e verbos de Hopper e Thompson (1984), é, como se viu, mais que uma abordagem de ordem semântico-cognitiva; é uma abordagem de ordem pragmática. Hopper e Thompson (1984) invocam argumentos discursivos para assinalar que o estatuto de um item lexical em sua categoria gramatical respectiva não é, de modo

algum, uma propriedade fixa da palavra em questão. As propriedades semânticas relevantes podem variar de acordo com o contexto de uso, variam na medida em que nomes se referem ou não a uma coisa identificável, estável, e verbos se referem ou não a um evento dinâmico específico.

Verifiquemos como essas propriedades se aplicam do mesmo modo ao português. Considerem-se os exemplos contidos em (6-13).

(6-13) a *então... o nazismo...* *<u>matou</u>:: <u>um:: cacetada de:: judeus...</u>* (NURC-SP-D2-343:53)
b *<u>Matar judeu</u> foi uma escolha consciente dos nazistas.*

Em (6-13a), tanto o nome *judeus* quanto o verbo *matou* são usados como membros prototípicos de suas respectivas classes. Em consonância com essa função referencial, tanto o verbo quanto o nome podem assumir o conjunto total de determinações e modificações que tipicamente recebem. O verbo, por exemplo, recebe marca de tempo, que é, nesse caso, o pretérito, aspecto simples, não o progressivo e modo indicativo e não o subjuntivo e voz ativa, não a passiva.

De modo similar, o nome pode assumir os vários atributos de sua classe, como aparecer no singular ou no plural, ser precedido por determinantes, ser modificado por adjetivos e orações relativas. Já no exemplo construído de (6-13b) nem o verbo *matar* nem o nome *judeu* exercem sua função referencial prototípica; em razão disso, o verbo *matar* aparece na forma de infinitivo e o nome, como um item lexical nu.

O processo de referenciação coloca o nome em distribuição complementar com o verbo. Um membro prototípico da classe dos nomes é o que denota um objeto visível, tangível, e um membro prototípico da classe dos verbos é o que denota uma ação efetiva, visível, cinética e concreta, envolvendo participantes e por eles executada. É o que, para Givón (1984), são formas de *realis*, isto é, predicados verbais que representam estados de coisas factuais, em

geral representados por orações principais declarativas no presente ou no pretérito

Uma causa da descategorização de verbos é representada por formas de *irrealis*, ou não factuais (Givón, 1984; Hopper; Thompson, 1984). As formas de *irrealis* são especialmente encontradas em orações contrafactuais, e outras orações não factuais que expressam volição, desejo ou ordem como (6-14), em que o complemento do volitivo se traduz pelo subjuntivo e o principal de um modal deôntico, como *ter que*, fica na forma não prototípica de infinitivo, não por acidente denominado "forma nominal do verbo" pela gramática tradicional.

(6-14) *e a menina porque quer que seja a (mamãe) que faça né? então sou eu que:: tenho que ir fazer et cetera et cetera* (NURC-SP-D2-360)

No português, o subjuntivo e o imperativo são destituídos de marcas verdadeiras de tempo e aspecto: os tempos assim chamados imperfeito e futuro do subjuntivo servem apenas para assinalar diferentes graus de não factualidade, com um grau razoável de possibilidade de ocorrência do evento no futuro (6-15a), e um grau mínimo de possibilidade de realização, no imperfeito simples (6-15b) e um grau zero de realização no imperfeito composto (6-15c).

(6-15) a *quando o menor... já:: estiver assim... pela quarta série terceira quarta série... ele já estará mais... independente* (NURC-SP-D2-360)
b *se o menor... já:: estivesse assim... pela quarta série terceira quarta série... ele já estaria mais... independente.*
c *Se o menor... já:: tivesse estado assim... pela quarta série terceira quarta série... ele já teria estado mais... independente*

Quanto maior o grau de dependência de um verbo a outro, menor o grau de categorialidade e, portanto, quanto maior o grau de dependência, tanto menor o grau de prototipicidade (Givón, 1981); um caso extremo são os complementos de causativos como *fazer*, conti-

do em (6-16a); numa escala abaixo, estão os complementos do tipo *que + subjuntivo* de verbos como *ordenar* e *insistir, esperar*; como em (6-16b). Já a forma finita do complemento de verbos como *achar*, em (6-16c), cujo grau de dependência em relação à matriz é menor, não se distingue de um membro prototípico da classe dos verbos.

(6-16) a *seria um um um bom... um bom jeito de de de se fazer chegar ao público as peças teatrais que estão... em cartaz* (NURC-SP-DID-234)

b *então... esperamos que não:: haja maiores problemas* (NURC-SP-D2-360)

c *aí eu achei que realmente a escola não preenchia tudo... que eu gostaria (que) preenchesse então eu tirei* (NURC-SP-D2-360)

Hopper e Thompson (1984) referem-se à natureza dependente da construção encaixada, para acrescentar um argumento discursivo: os falantes destacam eventos relatados, que são colocados estrategicamente como parte da figura no discurso em progresso, sendo, assim, devidamente distinguidos de material avaliativo e descritivo, que constitui o fundo da narrativa.

O falante pode também incluir, como pressupostos discursivos, eventos não abertos à discussão e que, por isso, são subsidiários dos eventos relatados, que são destacados como figura. Incluem-se nominalizações, orações encaixadas, principalmente a final, que exigem outra forma nominal, como infinitivo, (6-17a) orações relativas com formas nominais do verbo, como gerúndio (6-17b) e particípio (6-17c).

(6-17) a *estudei bem fiz um estudo certinho para ver qual era a melhor e foi determinado... foi visto que aquela era melhor* (NURC-SP-D2-360)

b *vão encontrar em muitos lugares... aqui no norte da Espanha... e depois... ( )... no sudoeste da França... e ainda aqui em toda essa região abrangendo o sudoeste da França* (NURC-SP-EF-405)

c *"Roda Viva" você assistiu?... palavrões escritos falados e distribuídos ((ri)) de todo jeito ((risos)) mas foi bom também viu?* (NURC-SP-DID-245)

É possível referir-se ao predicado verbal com seus modificadores como uma entidade do discurso, quando ele for estendido por operadores que representam o grau de factualidade (Givón, 1981). Somente quando um estado de coisas descrito numa predicação de fato ocorreu ou está ocorrendo, isto é, constitui um evento *realis*, é que ele pode ser temporalmente localizável e devidamente marcado com morfemas modo-temporais em português.

Eventos *irrealis* não podem ser temporalmente marcados porque (i) a predicação é expressa como uma ordem, isto é, sua realização é atribuível à ação do ouvinte no futuro; (ii) a ocorrência possível do evento denotado pela predicação é objeto de esperança, de temor ou de dúvida; (iii) ou o falante não tem evidência direta da realização do evento relatado.

Em uma oração dependente, ou construção encaixada, em que verbos atuam como nominalizações, o predicado passa a ser argumento de outro verbo. Segundo Hopper e Thompson (1984), essa é a situação suficientemente óbvia em que um estado de coisas é apresentado como pressuposto e, portanto, não submetido a condições de contra-argumentação.

Como visto anteriormente, o estudo transistêmico de Mackenzie (1987; 1996) mostra que uma construção encaixada pode estar ordenada em uma hierarquia entre "completamente verbal" e "completamente nominal". Em um extremo, o grau 0, a construção encaixada pode ser uma construção completamente finita, plenamente verbal. No outro extremo, grau 4, a construção encaixada, na forma de uma nominalização, adquire propriedades características de termos nominais, como operadores de termos, distinção de gênero e de número e modificadores adjetivais.

Essa hierarquia tem muito a ver com a discussão sobre nominidade prototípica. Com efeito, nominalizações não constituem membros centrais nem da categoria dos nomes nem da categoria

dos verbos; geralmente representam uma entidade de segunda ordem, ou um estado de coisas localizável no tempo e no espaço, como os verbos, mas, ao mesmo tempo, recebem atributos morfossintáticos próprios do nome. As nominalizações que podem ser consideradas instâncias prototípicas são as de primeira ordem, conforme já mostrado com o uso do termo *construção* em (6-6b). Outro caso ilustrativo é o do nome *desenho*s em (6-18), que, além de se referir a uma entidade concreta, é também discursivamente manipulável, nos termos de Hopper e Thompson (1984).

(6-18) *em alguns dos desenhos das cavernas principalmente em Altamira... há uma fidelidade... linear à natureza* (EF-SP-405:57)

Segundo Neves (1993; ver também Camacho; Dall'Aglio-Hattnher; Gonçalves, 2008), como classe de palavras o nome tem um estatuto categorial e um estatuto relacional. O estatuto categorial é observável em suas funções de denominação e descrição do denominado; o relacional é observável em sua função de referenciação que constitui o suporte para o cumprimento do complexo de funções que essa parte do discurso exerce no texto. Numa visão já clássica, Lyons (1977) distingue nomes, como *carro*, que têm uma denotação, de nominais, ou SNs, como *aquele carro vermelho*, que têm uma referência e, como tal, podem introduzir ou manter uma entidade no discurso, nos termos de Hopper e Thompson (1984).

Em sua função denominadora o nome faz uma designação de entidades cognitivas e/ou culturais que apresentam certas propriedades categorizáveis no mundo extralinguístico. Nos termos de Hopper e Thompson (1984), é somente na função referenciadora que os nomes exercem sua função prototípica como classe de palavra. Na função denominadora, essa prototipicidade fica suspensa.

Em sua função referenciadora, o significado do nome depende das relações com os outros elementos do SN, já que só se torna referenciador ao ser atualizado num SN. Ocorrer num SN é, entretanto, condição necessária, mas não suficiente, para que um nome exerça função de referência, já que em função predicativa, qualificadora,

o N já não referencia (Neves, 1993). Parece existir evidência de complementaridade e exclusão mútua entre as duas funções: uma basicamente substantiva/pronominal (a referenciadora), e outra, basicamente adjetiva (a qualificadora).

É relevante, a esse respeito, discutir o modo como a GDF enfoca a função denominadora e a referenciadora do nome, desenvolvendo unidades de análise em diferentes níveis de organização, especialmente os que distinguem funções pragmáticas de funções semânticas. Vale lembrar que o termo *semântica* se restringe aos significados de unidades lexicais e unidades complexas, independentemente do modo como sejam usadas na comunicação. O uso que se faz das unidades linguísticas é tratado no Nível Interpessoal, em termos de Atos e Subatos que especificam as funções das unidades linguísticas.

A relevância dessa distinção é particularmente aplicável à noção de referência. Numa sentença como *eu vi um leão,* há dois modos de se considerar a expressão *um leão* como uma expressão referencial: (i) é o falante que se refere a um animal da classe dos leões mediante o uso dessa expressão; (ii) é a expressão em si que se refere a um animal da classe dos leões. No caso de (i), toma-se a perspectiva interpessoal, acional e, no caso de (ii), a perspectiva representacional, semântica. Para distinguir esses dois usos da noção de referência, a GDF prefere o termo "referência" para a primeira interpretação e "denotação" para a segunda. É justamente na função de referência que o nome é mais prototipicamente categorial.

Feita essa distinção, é fácil verificar que uma expressão que denota um leão não é necessariamente empregada para se referir um leão, como no uso predicativo, quando se afirma que x é um leão; nesse caso, a expressão que denota um leão é usada atributivamente em vez de referencialmente: a denotação não muda, mas a função sim.

O termo-chave para descrever a natureza das unidades interpessoais é *evocação*, enquanto o termo introduzido para descrever a natureza de unidades representacionais é *denotação*. No Nível Representacional, uma expressão empregada para evocar uma entidade referencial é representada pela variável (x), enquanto uma

expressão usada para atribuir uma propriedade é representada pela variável (f).

O principal sintoma da perda de categorialidade por ausência de referência é o fato de que nenhuma das duas palavras pode ser flexionada ou receber modificadores. Esse fenômeno geral é considerado por Hopper e Thompson (1984) como um universal linguístico. Os nomes podem referir-se potencialmente a uma entidade discreta no Nível Representacional, mas nem sempre exercem função referenciadora no contexto em que aparecem no Nível Interpessoal. Observe as sentenças contidas em (6-19) e (6-20).

(6-19) a *agora não é simples você usar computador...* (D2-SP-343)
b *Você usou o computador azul do laboratório?*
c *Computadores são meros instrumentos de trabalho.*

(6-20) a *tirou diploma com::... com menção honrosa* (D2-SP-360)
b *Tirou da gaveta o meu diploma de medicina para fazer uma cópia.*

Os SNs grifados das sentenças de (6-19a) e (6-20a) ilustram a parcial descategorização dos nomes nucleares, em função de um uso não referenciador e, portanto, não discursivamente manipulável, de *computador* em (6-19a) e de incorporação de *diploma* a um verbo-suporte em (6-20a).

Os nomes nesses casos não podem ser flexionados, ser modificados, nem determinados, exatamente o que ocorre com o uso dos mesmos itens lexicais em (6-19b) e (6-20b), que consistem em exemplos altamente representativos da classe de nomes. Também *computadores* de (6-19c) não se refere a uma entidade discreta e identificável, mas toma a entidade referida como uma classe genérica.

Essa variabilidade nas propriedades do nome ao longo do discurso sugere que qualquer divisão absoluta, não contextual do léxico, em raízes verbais e nominais, terá uma validade somente limitada. O fato semântico de que um radical denota uma entidade visível não é crucial para torná-lo um nome prototípico. Em vez

disso, é preciso que ele exerça certo papel no discurso em que figura, de modo que a nominidade prototípica serve para introduzir um participante, o que implica uma identidade específica do referente, não uma entidade genérica, como em (6-19c).

O grau máximo de prototipicidade é de fato atingido quando alguma entidade individuada é introduzida como participante no discurso, como em (6-21a-b), com paralelismo perfeito entre Referência (Nível Interpessoal) e denotação (Nível Representacional).

(6-21) a *Olha eu estive:: o mês passado em Poços de Caldas estava passando um filme que eu achei lindo aqui em São Paulo A filha de (Ryan) é um filme... éh muito bom de após guerra é um filme de amor... umas cenas maravilhosas... lindo o filme... eu assisti faz tempo já* (NURC-DID-SP-234: 112)

b *eu tenho assistido umas PEças eu assisti u::ma com a:: aquela artista magrinha de televisao aquela moreninha que é bailarina também*... (NURC-DID-SP-234)

Observe-se em (6-21a) que o nome *filme* é núcleo de um SN que apresenta uma entidade referencial pela primeira vez no discurso corrente, e em (6-21b), que o falante se empenha em construir um referente já disponível, presumivelmente, na memória de longo prazo do ouvinte; nesse caso, o nome nuclear (*artista*) é regularmente acompanhado no SN por determinação (demonstrativo *aquela*), modificação (adjetivo *magrinha*) e localização (sintagma preposicional *de televisão*).

O traço de maior grau de relevância linguística é o da saliência discursiva. Formas, como *filme* em (6-21a) e *artista* em (6-21b), que exercem um papel no discurso – dado serem manipuláveis ou desenvolvíveis nele – são universalmente capazes de adotar o conjunto total de atributos gramaticais de nomes: determinantes, afixos de caso (quando for o caso), gênero e número, bem como classificadores, modificadores e, em certas línguas, morfemas assinalando nominidade. Esses nomes representam entidades autônomas, usualmente concretas e individuadas; podem atuar como partici-

pantes e suportes do discurso e podem ser referidos indiretamente mediante processos anafóricos. Esse processo de referenciação tem uma inter-relação extremamente interessante com outros processos linguísticos como a dêixis e a foricidade.

É possível distinguir estratégias que descategorizam nomes com o estatuto de informação dada. Uma classe dêitica que interage fortemente com a dos nomes em função referenciadora é a dos pronomes. No caso dos pronomes de 1ª e 2ª pessoas, que representam a correlação de pessoa (Benveniste, 1976), a ausência de nomes referenciais em proveito dessa categoria dêitica é simplesmente uma regra categórica.

Já quanto ao uso dos pronomes referentes à não pessoa, como os de 3ª, a situação é crucialmente distinta. Ao substituir o nome, o uso de pronomes, uma estratégia discursiva para referência à informação dada, representa perda de especificidade semântica e apenas retenção de categorias semânticas e discursivas mais gerais, como caso, gênero, número e dêixis.

O enunciado contido em (6-22) é um bom exemplo dessa interação entre processos de referência, dêixis e foricidade, mediante uma curiosa interação entre as diferentes classes de palavras.

(6-22) *depois eh:: terça e quinta... <u>a menina</u> faz fonoaudiologia porque <u>ela</u> está com três anos e pouco... e ainda não ∅ fala... ∅ fala muito pouco... então <u>ela</u> faz... reeduca/... reeducação não mas seria... exercícios... com a fonoaudióloga para ver se::... se ∅ começa a falar mais rapidamente...* (D2-SP-360)

O nome manipulável *menina* perde parte de sua especificidade semântica ao ser mencionado no discurso corrente pelo anafórico *ela*, que retém marcação do caso reto, de gênero e de número. Num estágio mais avançado, essa estratégia de referenciação ao constituinte discursivamente mais saliente e, portanto, mais recorrente no discurso em progresso faz com que o pronome anafórico seja simplesmente elidido. Esse procedimento significa retomar no texto um referente discursivamente saliente mediante o recurso à

anáfora zero, que é capaz de recuperar, desse modo, as funções sintáticas do nome, como a de sujeito do verbo *input* no caso de (6-22).

Há nomes que se referem a uma informação já conhecida, como *cenas* em (6-23) e a primeira menção de *crianças* em (6-24). Esse tipo de nome ocupa uma posição intermediária entre formas prototípicas referindo-se a entidades apresentadas pela primeira vez no discurso, como (6-21a-b) acima, e formas não manipuláveis representando uma classe de entidade e não exatamente uma instância referencial específica e individualizada, como a segunda menção do lexema *criança* em (6-24).

(6-23) *outro filme que me chocou bastante também... dada as cenas brutas de de de m/ de mata e mata que matou um mundo de gente foi "Bonnie and Clide" eu acho que foi... não lembro faz tempo que assisti ∅ também mas aquele filme saí CANSAda do cinema cinema* (DID-SP-234)

(6-24) *eu trago muito processo para casa e faço em casa... que para eu voltar lá à tarde cria muito problema com as crianças* [...] *e as coisas de casa que a gente aten/ tem que atender normalmente com crianças BRIgas que a gente tem que repartir* (D2-SP-360)

Em termos da hipótese da categorialidade de Hopper e Thompson (1984), nomes que se referem a uma entidade informacionalmente dada e inferível podem identificar-se como os que se referem a entidades novas, como o SN *as cenas brutas* de (6-23), que é dotado de determinação e modificação.

Por outro lado, nomes que se referem a uma informação dada podem, alternativamente, assumir a morfologia de nomes genéricos, não específicos, como a segunda menção de *crianças* em (6-24): esse SN estabelece uma relação anafórica com a primeira menção, mas não num valor discursivo manipulável, já que, hiperonimicamente, trata-se agora simplesmente do conjunto que contém as instâncias reais da entidade anteriormente mencionada. Além disso, como em todas as línguas examinadas por Hopper e Thomp-

son (1984), também em português é possível se perder o conteúdo semântico da retomada anafórica de um nome, retendo-se apenas o resíduo discursivo-gramatical.

Desse modo, esquematicamente, é possível aplicar ao português o mesmo *continuum* categorial contido em (6-25).

(6-25)

CATEGORIALIDADE ALTA:	formas "apresentacionais"
CATEGORIALIDADE INTERMEDIÁRIA:	formas anafóricas/contextualmente estabelecidas
CATEGORIALIDADE BAIXA:	proformas, formas genéricas/não referenciais, morfemas de concordância, anáfora zero

(Adaptado de Hopper; Thompson, 1984, p.723)

A falta de homogeneidade dos membros da terceira categoria sugere para Hopper e Thompson (1984) o estatuto de não marcação, posição que fornece apoio para o postulado de ser a anáfora zero o grau mais baixo de categorialidade.

6.3 Paralelismo formal entre a classe referenciadora e a classe predicadora

É inegável a contribuição das duas classes maiores – o nome e o verbo – e as duas intermediárias – o adjetivo e o advérbio – para a construção da sentença, seja na função de referenciação, seja na função de predicação. Na função referenciadora, sobressaem o nome e o adjetivo como principais entidades do SN, na medida em que constituem os dois principais restritores do SN (Dik, 1989). Na função predicadora, sobressaem o verbo e o advérbio. É verdade que adjetivos também são classes predicadoras, se for considerada apenas sua atuação no nível do SN, âmbito em que predica o nome.

Se considerarmos a organização em níveis e camadas para a representação da sentença, proposta por vários autores da GF e da GDF, mas mais sistematicamente por Dik (1989) e por Hengeveld

(1989; 2004a), torna-se inegável a existência de um paralelismo entre nomes e adjetivos, por um lado, e verbos e advérbios, por outro. É necessário, todavia, nos restringirmos aos membros prototípicos da classe dos advérbios, que são os advérbios de modo e velocidade, rotulados por Ilari et al. (2002) de advérbios "qualitativos" como *furiosamente* em *brigar furiosamente*; e também aos membros prototípicos da classe dos adjetivos, os que designam em geral uma propriedade mais ou menos inerente do nome.

Como também o nome é aqui considerado em relação a entidades (espaciais) de primeira ordem, tais propriedades podem envolver domínios como tamanho, cor, peso etc., o que produziria um SN simples como *gato preto*. Considere-se a sentença potencial em (6-26).

(6-26) *Aqueles dois gatos pretos no tapete da sala brigaram furiosamente duas vezes nesta manhã.*

Observa-se que às operações de qualificação do nome (*pretos*) e do verbo (*furiosamente*) sobrepõem-se duas operações, a de quantificação e a de localização. No nome, a quantificação é operada pelo uso do numeral *dois*, e a de localização pelo sintagma preposicional *no tapete*; no verbo, a quantificação é operada pelo adverbial *duas vezes* e a localização pelo adverbial de tempo *nesta manhã*.

No SN *os dois gatos pretos no tapete*, às entidades referidas por *os gatos* se aplica a propriedade "ser preto", do que se excluem a quantidade e o locativo. O modificador quantitativo *dois* especifica o número das entidades, não o número de localizações; tanto *aqueles* quanto *no tapete da sala* especificam a localização das entidades "gato" com as propriedades qualitativas ("ser preto") e quantitativas ("ser dois") fornecidas no SN.

Quanto ao verbo, as relações de escopo entre as propriedades qualificadora (*furiosamente*), quantificadora (*duas vezes*) e a localizadora (*nesta manhã*) são exatamente idênticas às relações de escopo das mesmas propriedades do nome. As operações de qualificação, quantificação e localização podem manifestar-se mediante o uso de morfemas gramaticais ou de morfemas lexicais, conforme se ob-

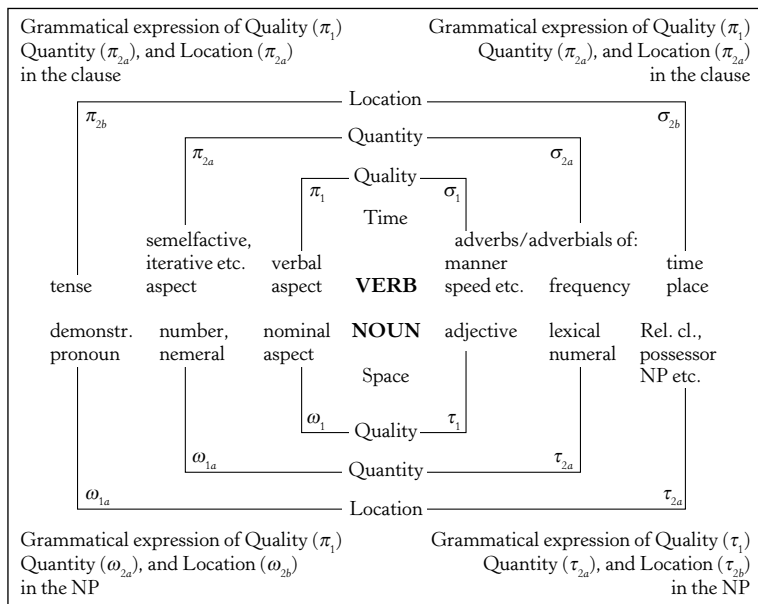
serva na Figura 11, que mostra o paralelismo entre a representação subjacente do nome e a do verbo.

(Rijkhoff, 2002, p.224)

Figura 11 – Simetria entre a representação subjacente da oração e a do SN

As indicações dêiticas, operadas no nome pelo artigo e no verbo pelas flexões modo-temporal e número-pessoal, são, como se vê, efetuadas por operadores gramaticais, enquanto as outras atribuições são feitas por recurso ao léxico, mediante o uso de advérbios e de adverbiais.

A explicação mais provável para essa simetria pode estar no fato de que, como entidades temporais, os verbos podem ser entendidos, na esteira de Givón (1979), como termos de entidades espaciais cognitivamente menos complexas que o nome; assim, o modo como o falante constrói a referência a estados de coisas é modelado de acordo com o modo como constrói a referência a SNs (Rijkhoff, 2002).

Para Dik (1989), são as regras de expressão que finalmente dispõem a representação subjacente do SN e da oração em uma forma ordenada para ser finalmente enunciada; para Hengeveld e Mackenzie (2008), a operação de codificação é que dá formato morfossintático e fonológico para as representações pragmáticas e semânticas dos níveis superiores de formulação.

Assim, é pertinente observar que as regras de expressão morfossintática do português dão preferência por alocar no lado esquerdo do núcleo do SN a expressão gramatical de propriedades pragmáticas, reservando para o lado direito do núcleo a expressão semântica (cf. Camacho; Dall'Aglio-Hattnher; Gonçalves, 2008). Assim, a determinação e a quantificação costumam ser acompanhadas por adjetivos menos prototípicos como *mero*, *pobre*, *simples* etc. Essa subclasse de adjetivos não atua como modificadores do nome, como os membros prototípicos da classe, mas refletem uma atitude do locutor no discurso corrente. Outros subtipos como *antigo* situam a permanência da referência numa situação cronologicamente anterior ao momento da fala, manifestando também propriedades nitidamente pragmáticas.

É verdade que esses modificadores discursivos aparecem apenas nos SNs que atuam como predicadores mediados por cópula, em que o nome não exerce sua função referenciadora prototípica, como (6-27a), e são estranhos em SN manipuláveis representando entidades de primeira ordem, como em (6-27c).

(6-27) a *a imagem é uma [simples, reles, pobre] composição de motivos* (NURC-SP-EF-405)

b *Aquela reles/pobre/simples composição de motivos não me dá a impressão de ser uma arte verdadeira.*

c (?) Aquele simples/reles/professor de linguística que mora na esquina aposentou-se o ano passado / José é um simples/reles/pobre professor de linguística.

Esses aspectos discursivos lembram a necessidade de estabelecer uma diferenciação entre a expressão linguística do SN, com

seus *modificadores descritivos* de qualificação, quantificação e localização, e o SN como uma entidade referenciadora na instância discursiva em que é de fato empregado, ou seja, o SN em face de seus *modificadores discursivos*.[7] Os artigos constituem operadores discursivos já que permitem ao falante determinar se os referentes são identificáveis (definidos) e não identificáveis (indefinidos) para o ouvinte, isto é, se são manipuláveis no discurso, nos termos de Hopper e Thompson (1984).[8]

Assim, os SNs que funcionam como objetos espacialmente delimitados, os de primeira ordem, têm uma função dupla: são, por um lado, sintagmas descritivos que podem simplesmente exercer a função de descrição do denominado, mediante a especificação de propriedades qualitativa, quantitativa e locativa de que dispõe a entidade denotada; constituem, por outro, expressões referenciadoras, manipuláveis, que servem de suporte para o cumprimento do complexo de funções que os nomes exercem no texto, na medida em que apontam para alguma entidade discursivamente saliente.

Em vista da dualidade estrutural entre o SN e a oração, é possível estender, segundo Rijkhoff (2002), a simetria funcional entre SNs e orações, já que, como o SN, a oração pode tanto fornecer uma descrição de uma entidade quanto referir-se a essa entidade no discurso. Dito em outros termos, é possível traçar, no nível da oração, cujo predicado nuclear é o verbo, uma distinção entre a descrição que o falante faz de uma entidade, no caso a estrutura linguística da oração, e o referente da descrição do falante, ou seja, o próprio evento.[9]

7 Vale recordar aqui a distinção fregeana de base formal entre sentido e referência, que permite operar uma interpretação não tautológica da sentença "A Estrela da Manhã é a Estrela da Tarde" (cf. Lyons, 1977).

8 Outros mecanismos lexicais que estabelecem relações discursivas são satélites como *acima/antes mencionado,* que especificam quando e onde o referente foi introduzido no discurso prévio.

9 Há razões de referenciação anafórica para essa distinção, tanto no nível do SN quanto no nível da oração. Sempre é possível fazer referência tanto ao conteúdo quanto à forma de expressão (ver exemplos (2-16) a (2-19) na Seção 2.4). É essa relação que está por trás da distinção proposta por Rijkhoff (2002).

Desse modo, os operadores factual e não factual da oração se correlacionam com os operadores definido e indefinido do SN, já que relacionam o evento como uma espécie de referente discursivo; os operadores de factualidade e definitude têm a função similar de sinalizar que os referentes dos respectivos itens lexicais, eventos e entidades já existem no mundo do discurso ou sua existência é no mínimo pressuposta; por outro lado, o que suas contrapartes negativas, não factualidade e indefinitude, têm em comum é que as entidades a que elas se referem não existem ou, pelos menos, ainda não têm existência no mundo do discurso como entidades identificáveis ou factuais (cf. Rijkhoff, 2002).

Pode-se acrescentar que, em outros termos, os operadores discursivos de factualidade e de definitude permitem localizar ocorrências de verbos e de nomes e seus respectivos modificadores como entidades prototípicas dos processos de predicação e de Referência.

6.4 Classes de palavras e categorização tipológica do português

Retomando as quatro classes de predicados discutidas na Seção 6.3, interessa-me agora vê-las sob uma perspectiva tipológica em relação aos sistemas de partes do discurso elaborada por Hengeveld (1992), conforme se vê na Figura 12.

Generalizando observações como essas, Hengeveld (1992) assinala que os sistemas de classes de palavras podem ser divididos em dois grupos principais de línguas *diferenciadas*, como o inglês e o português, que têm quatro diferentes classes de palavras exercendo quatro diferentes funções, e de línguas *não diferenciadas*, como o holandês, com menos de quatro das categorias de predicados presentes nas línguas diferenciadas.

Generalizando ainda mais essas observações, sistemas *não diferenciados* de classes de palavras podem ser subdivididos em dois outros grupos: o das línguas *flexíveis*, em que uma única parte do discurso pode ser usada com diferentes funções, como o holandês;

	1	V/N/A/Adv				tongan
Flexível	2	V	N/A/Adv			quechua
	3	V	N	A/Adv		holandês
Diferenciada	4	V	N	A	Adv	inglês
	5	V	N	A	---	wambon
Rígida	6	V	N	---	---	!xù[11]
	7	V	---	---	---	tuscarora

(Adaptado de Hengeveld, 1992, p.69)

Figura 12 – Sistemas de classes de palavras

e o das línguas *rígidas*, em que falta uma parte do discurso para o exercício de algumas funções.

Um caso de língua com grau máximo de flexibilidade é o tongan (língua polinésia falada nas ilhas de Tonga), para a qual uma única classe pode ser empregada, sem qualquer modificação formal, para traduzir todas as funções predicadoras exercidas por verbos, nomes, adjetivos e advérbios no português.

Um caso extremo de língua rígida é o tuscarora (língua da América do Norte da família do iroquês), que dispõe de apenas uma classe de palavras, o verbo; desse modo, para traduzir, por exemplo, o significado de nomes típicos do português é necessário usar uma predicação verbal.

O português é uma língua diferenciada, já que dispõe de uma classe de palavras para cada uma das quatro principais funções predicadoras, atributo que a situa na região média do *continuum* contido na Figura 12. Esse *continuum* começa com línguas extremamente flexíveis como o tongan, em que a mesma categoria lexical exerce, sem adaptação morfológica, qualquer função predicadora, e termina com línguas extremamente rígidas como o tuscarora, que tem que adaptar a uma predicação verbal todos os outros três tipos de predicação.

10 !xù é uma língua da família Khoisan (Coisã) falada no sudoeste da África, principalmente Botsuana e Namíbia, pelos bosquímanos ou hotentotes.

Para dar um exemplo simples, o tuscarora necessita fazer uma predicação do tipo de "ele é jovem" para traduzir o que o português denota com o nome *menino*. Assim, para traduzir o que se diria em português com *o menino olhou para o bode*, seria necessário criar em tuscarora aproximadamente uma sequência de predicações do tipo "ele é jovem, ele olha para ele, ele fede".

Qual seria, então, o correlato gramatical mais evidente para as classes de palavras desse traço tipológico do português? Como língua diferenciada, nem flexível, nem rígida, a gramática do português necessita, em geral, de ajustes formais para converter um item lexical de uma classe de predicados num item lexical de outra. A morfologia derivacional do português é rica em afixos, conforme mostra o seguinte exemplo: *forma* [N] > *formar, deformar, reformar* [V] > *formoso, formal, deformado, reformável* [Adj] > *formosamente, formalmente* [Adv].

Há, por outro lado, casos de flutuação categorial entre formas fonologicamente idênticas, mas com distintos significados relacionados e distintas funções e/ou propriedades sintáticas, como a que ocorre também entre predicados lexicais adjetivais e adverbiais, conforme se observa em (6-28).

(6-28) *então eu expliquei <u>direitinho</u> que se realmente for bom vocação eu:: não impedirei de seguir* (D2-SP-360)

Basílio (2002) considera que a transposição de adjetivo para advérbio, como se vê em (6-28), implica o abandono de propriedades sintáticas anteriores, como a de modificador adjetival, e a preservação de propriedades semânticas no exercício de função de outra classe, no caso, a de advérbio. Esse fenômeno, tratado como um processo por conversão, é similar ao processo de derivação lexical que ocorre com a palavra *policial,* que tanto pode ser um adjetivo quanto um nome.[11] O fato é que a necessidade de conversão, com

11 Esse fenômeno é diferente do de "extensão de propriedades", em que um item lexical ou classe de itens assume apenas uma ou outra das propriedades de uma outra classe como o adjetivo *aflito* e o nome *os aflitos* (Basílio, 2002, p. 75-76).

ou sem forma fonológica idêntica, é um traço distintivo de línguas diferenciadas, como o português.

As nominalizações deverbais constituem exemplos flagrantes de predicados verbais que necessitam submeter-se a ajustes formais próprios de um modelo prototípico de termo, que é o nome denotando entidades de primeira ordem, ou seja, os chamados nomes concretos (Dik, 1985; Lyons, 1977).

Vale recordar aqui os ajustes mais comuns da predicação verbal encaixada à expressão do termo nominal, já mencionados na Seção 3.2: um predicado verbal transforma-se num núcleo nominal; um operador de predicado verbal, como o sufixo modo-temporal, vira zero no predicado nominal e, inversamente, um zero no predicado verbal, como a noção de definitude expressa pelo artigo, transforma-se num operador de termo no predicado nominal; o primeiro e o segundo argumento podem tanto assumir a forma de uma expressão de possuidor quanto a de um adjetivo; já um satélite adverbial só pode assumir a forma de um adjetivo (Dik, 1985; 1997). Os predicados verbais de (6-29a) e nominal de (6-29b) são casos ilustrativos desses ajustes.

(6-29) a *O Brasil comprou rapidamente os dólares do mercado.*
b *A rápida compra brasileira dos dólares do mercado.*

O mesmo é verdadeiro para os predicados nominais derivados de predicados adjetivais, os quais, na condição de monoargumentais, tendem simplesmente, como os nominais derivados de verbais de um lugar, a preservar o argumento central, como se observa em (6-30a-b).

(6-30) a *eles conseguem chegar a uma fidelidade linear... da natureza... à extrema exatidão do desenho* (EF-SP-405).
b *A natureza é linearmente fiel e o desenho é extremamente exato.*

Observem-se, além disso, outras mudanças categoriais, mediante o uso de processos derivacionais, como as que ocorrem entre *linear* e *linearmente* e *extrema* e *extremamente*. Esses ajustes permi-

tem limitar os tipos de nominalização que podem variar entre quase completamente verbais para quase completamente nominais.

Uma boa pergunta a fazer aqui é a seguinte: dado que os predicados deadjetivais e deverbais não constituem membros prototípicos de suas respectivas classes, por que uma língua diferenciada necessita de processos de mudança categorial, como os que se dão com a derivação, cujo resultado contradiz justamente sua natureza diferenciada centrada nos membros mais prototípicos?

A resposta está justamente de novo na relação entre as categorias e o discurso. Por situarem-se no conjunto das predicações encaixadas e, portanto, atuarem como termos, os nomes derivados são dotados de uma especificidade gramatical própria: além de preservarem praticamente intacta a estrutura valencial do termo primitivo que lhes deu origem, podem funcionar como termo na predicação matriz, assumindo funções sintáticas, semânticas e pragmáticas que, de outro modo, não seria possível assumir, caso se mantivessem como verbos.

Em termos informacionais, a apresentação de uma predicação como núcleo de um SN é um convite ao interlocutor para atentar para alguma entidade referencial. No caso de uma nominalização, o referente não pode ser uma pessoa, animal ou outro objeto tangível, deve ser uma entidade de ordem superior, ou seja, de segunda ou de terceira ordem.[12] A entidade selecionada deve fornecer apenas a informação relevante; saber, por exemplo, que é suficiente nomear o predicado dispensa a especificação formal de todos os argumentos, como se vê no nome grifado de (6-31).

(6-31) *mas é possível a gente olhar para elas e ainda se espantar com a QUAlidade da <u>representação</u> então são dois fatos diferentes...* (EF-SP-405)

---

12 Nomes referentes a entidades de segunda ordem representam um estado de coisas designado pela predicação encaixada, enquanto nomes referentes a entidades de terceira ordem representam uma proposição designada pela predicação encaixada ou seu correlato intensional (Lyons, 1977).

Em termos sintáticos, uma oração subordinada finita, que costuma aparecer no início da sentença, pode, por razões de complexidade estrutural, ser fortemente motivada a adotar a posição final. Como construção encaixada, uma expressão nominalizada dispõe, contudo, de menor complexidade categorial e, consequentemente, de maior versatilidade sintática, podendo, por isso, aparecer na posição de sujeito, de objeto, ser regida por preposições etc. A sentença (6-32a), por exemplo, licencia a aplicação de uma predicação avaliativa ao sujeito *influência indígena sobre a alimentação*; entretanto, se o estado de coisas fosse expresso por um predicado verbal, como em (6-32b), o resultado seria uma construção agramatical.

(6-32) a *A influência indígena sobre a alimentação é muito grande.*
b (?) *Os indígenas influírem sobre a alimentação é muito grande.*

De um ponto de vista textual, favorece a escolha de uma nominalização a introdução explícita de um novo referente discursivo. A nominalização permite nomear a entidade de ordem superior, um nome que desfruta de uma versatilidade de uso que complementa a versatilidade sintática e informacional acima mencionada (cf. Camacho; Santana, 2004). (6-33) contém nominalizações com diferentes estatutos informacionais, como nova não usada (6-33a), inferível (6-33b) e textualmente evocada (6-33c) (Prince, 1981).[13]

(6-33) a *eles viviam basicamente da <u>coleta</u> eram caçadores...* (EF-SP-405)
b *o paleolítico é período período... da pedra lascada... como vocês todos sabem... não é?... e... tem uma <u>duração</u> de aproximadamente seiscentos mil anos* (EF-SP-405).

13 De acordo com a taxonomia proposta por Prince (1981), uma entidade *nova não usada* é aquela em que o ouvinte supõe ter uma entidade correspondente em seu modelo discursivo. Já o referente de uma entidade *inferível* pode ser identificado pelo ouvinte, via razões lógicas ou plausíveis, a partir de outra entidade discursiva evocada ou inferível. Por fim, uma entidade *textualmente evocada*, ou simplesmente *evocada*, é indicada por uma entidade que já ocorreu no modelo discursivo do ouvinte e pode ser evocada no campo textual.

c *então a arte vai nascer:: em função dessa NEcessidade... de se manter vivo... necessidade que vai se caracterizar de forma PRINcipal:: em termos de comida...* (EF-SP-405)

Da contiguidade pragmática dos três tipos de entidades informacionais acima exemplificadas – novo não usado, inferível e evocado –, depreende-se que uma nominalização dificilmente constitui informação completamente nova. Uma nominalização constitui uma estratégia de pressuposição e, como tal, exerce uma função textual típica de representar uma entidade referencial discursivamente já conhecida.

Considerações finais

É tempo de balanço. Vimos que há uma possibilidade exploratória potencialmente relevante de aproximação entre classes de palavras e processos de descategorização e recategorização, o que me levou a fixar como objetivo comprovar a hipótese de que a nominalização tem uma natureza complexa em um *continuum* categorial. Ela figura entre duas classes potencialmente universais por atuar na constituição do ato locutório de formulação de uma sentença, a do nome, mediante um Subato de Referência, e a do verbo, mediante um Subato de Atribuição.

A continuidade categorial parece uma propriedade indiscutível da linguagem, pelo menos na tradição funcionalista, onde é tratada como um verdadeiro universal linguístico, evidenciada nos diferentes pontos de corte que caracterizam tipos regulares de línguas naturais. Além de buscar evidência sistemática para a comprovação desse axioma, este trabalho debruçou-se obsessivamente, digamos assim, sobre a estrutura argumental da nominalização, num esforço concentrado por demonstrar que esse mesmo princípio universal é metodologicamente útil e teoricamente válido para postular relações intralinguísticas de continuidade categorial mesmo entre classes aparentemente discretas como nome e verbo.

A trajetória percorrida para a sustentação da hipótese da continuidade categorial requereu, ainda, a demonstração da validade de

outra hipótese secundária, mas relevante para os objetivos gerais. Com efeito, foi necessário passar pela comprovação da hipótese de preservação de valência, postulada por Dik (1985; 1997), segundo a qual a estrutura argumental é parte constitutiva da nominalização, o que exigiu esforço quantitativo adicional com as evidências empíricas disponíveis numa amostra extraída da variedade culta do português falado do Brasil.

Esse esforço quantitativo ligou as duas versões do modelo, a GF e a GDF, ao mostrar que a expressão da estrutura valencial das nominalizações é motivada por princípios semânticos e pragmáticos que determinam a especificação ou não dos argumentos no *output* da gramática, não por processos de apagamento imunes a princípios semânticos e pragmáticos disponíveis ao falante na formulação do enunciado.

Na realidade, a postulação do princípio de economia sintagmática (Haiman, 1985; Cristofaro, 2003), que se vincula fortemente à vocação funcionalista deste trabalho, permitiu comprovar que a não especificação da valência constitutiva, nos termos da GF de Dik (1985; 1997), é o resultado de motivações pragmáticas e semânticas, emanadas, respectivamente, do Nível Interpessoal e do Nível Representacional agindo sobre os mecanismos tipicamente formais do Nível Morfossintático, nos termos da nova arquitetura proposta por Hengeveld (2004a) e consolidada por Hengeveld; Mackenzie (2008).

Essa busca não teria êxito se a trajetória percorrida não utilizasse um atalho necessário, representado pelo princípio de prototipicidade. De fato, postular a existência de categorias intermediárias como a de nominalização – uma categoria complexa, ora representando um Subato de Atribuição, próprio do verbo, ora um subato de Referência, próprio do nome – implica necessariamente a existência de membros mais prototípicos de uma categoria.

A existência de estrutura argumental, que sinaliza um nome denotando entidades de ordem superior, permitiu aproximar a nominalização de membros não prototípicos da categoria dos verbos, em companhia de formas não finitas, enquanto a ausência de estrutura

argumental, que sinaliza um nome de primeira ordem, permitiu aproximá-lo de membros prototípicos da categoria dos nomes.

As nominalizações ocupam justamente essa região amorfa, indistinta, a meio caminho entre membros prototípicos do nome e membros prototípicos do verbo. Na posição em que se situa na gramática, a nominalização não constitui um indivíduo solitário, dado que desfruta da companhia de outras categorias igualmente híbridas, como as formas não finitas, enquadramento que a tradição gramatical vem reconhecendo sistematicamente ao denominar infinitivos, gerúndios e particípios como formas nominais do verbo.

Um critério semântico que permitiu aproximar e agrupar essas formas categorialmente complexas na classe das construções encaixadas está na natureza aberta desses predicados. Com efeito, além da possibilidade de constituir o núcleo da construção encaixada, traço em comum com as formas nominais, que Dik (1985, 1997) reconhece, postulei que a nominalização se realiza como predicação fechada ou aberta, conforme a natureza especificada ou não especificada da estrutura argumental, traço adicional que essa categoria compartilha com infinitivos, particípios e gerúndios. Ainda que Dik (1997) atribua esse traço somente a construções infinitivas e participiais, sustentei, fincando bases nas evidências empíricas, que é conveniente incluir as nominalizações também na categoria dos predicados com posições abertas.

A evidência que dá suporte a essa categorização está justamente na falta de especificação formal da valência no Nível Morfossintático, representada por dois tipos de zeros anafóricos. Esses dois tipos de anáfora comprovam a existência de motivações em competição (Du Bois, 1985) para a formatação final da gramática, que são, em última instância, semânticas e pragmáticas.

Embora constituam motivações originadas em diferentes níveis de organização do Componente Gramatical, os dois casos de anáfora, representando não especificação de argumentos realmente constituivos da estrutura argumental da construção encaixada, são governados pelo mesmo princípio funcional de economia, que, como um dos principais universais que governam a língua enquan-

to instrumento de comunicação, representa uma pressão para a simplificação máxima da expressão.

O principal resultado desse princípio organizador, que explica a tendência pela omissão de informação redundante ou recuperável no contexto, é tornar a estrutura linguística menos transparente, em contraposição ao princípio da motivação icônica, o que no fim mostra por que a emergência discursiva da estrutura morfossintática não é sincronicamente visível. Essa invisibilidade da emergência discursiva das construções dá a impressão equivocada de que a gramática, além de autônoma, é também abolutamente arbitrária e desvinculada das motivações que governam a situação discursiva.

Na verdade, na situação discursiva emerge o sujeito falante, aparentemente invisível para a teoria linguística, que, socialmente identificado por suas relações com o interlocutor, constitui o agente responsável por escolhas significativas que processam contínuas mudanças no sistema linguístico. É verdade que as motivações semânticas da não especificação de argumentos parecem impor-se sob risco de agramaticalidade, mas não passam de um subproduto de um princípio mais geral de economia.

O modo como as duas motivações operam, por atuação mais ou menos direta dos participantes da interação, tem um correlato cognitivo evidente no processo de implementação dinâmica na GDF. Essa correlação permite aplicar às hipóteses aqui desenvolvidas o princípio de adequação psicológica, já avançado por Dik (1989). O modo como a nominalização se processa por derivação em Dik (1985), morfologicamente governada a partir de um verbo que lhe serve de *input* no léxico, é típico de línguas diferenciadas (Hengeveld, 1992), aspecto que fornece também uma justificativa bem fundamentada para o princípio de adequação tipológica, defendido pela GF e pela versão mais recente, a GDF.

Por fim, um ponto que merece destaque neste balanço final é a questão metateórica da consistência possível entre os fenômenos observados e a perspectiva teórica adotada. Um saldo positivo claramente derivado do investimento metodológico é o afastamento voluntário do que Du Bois (1985) denomina "funcionalismo

transparente". Apesar de assumir o princípio de que os fatores sintáticos, considerados autônomos por concepções formalistas, constituem, na realidade, os resultados transparentes dos objetivos funcionais do falante, de modo algum postulei que a teoria da linguagem dispensa totalmente motivações arbitrárias para a codificação morfossintática (Du Bois, 1985).

Defendi, na realidade, um conceito relativo de autonomia do Nível Morfossintático. Reconheci, por um lado, a existência de fenômenos morfossintáticos não diretamente compatíveis com aspectos semânticos e pragmáticos em consonância com o que propõem os proponentes da GDF, além de outros funcionalistas como Givón (1979).

Admiti, por outro, a existência de uma relação de interdependência entre a formulação, restrita ao Nível Interpessoal e ao Representacional, e a codificação, restrita ao Nível Morfossintático, por postular tanto uma motivação pragmática quanto uma motivação semântica na representação como zero anafórico de referentes previstos na memória de curto prazo dos interlocutores.

A posição moderada que defendi pressupõe a existência de muitos fenômenos linguísticos que sustentam a existência até certo ponto autônoma do Nível Morfossintático, com categorias próprias, mas a explicação das propriedades formais desse nível deve necessariamente fazer referência a fatores explanatórios, alinhados aos níveis Interpessoal e Representacional. Esses dois níveis, juntamente com o Morfossintático e o Fonológico, constituem o componente gramatical na concepção da GDF.

Assim, em vez de eliminar um nível independente, formal e autônomo da organização estrutural da linguagem, considerei preferível admitir que o Morfossintático é um nível independente da gramática, assim como os outros, mas, em certos aspectos, funcionalmente motivado e cujas propriedades formais refletem os parâmetros explanatórios de natureza pragmática e semântica que motivam sua emergência discursiva, nos termos de Hopper (1987). Diferentemente desse autor, entretanto, tenho plena convicção de que nem tudo que é morfossintático emerge dos dois níveis mais altos.

Em função desse posicionamento, procurei fornecer representações formais adequadas que pudessem sinalizar, nos níveis Interpessoal, Representacional e Morfossintático da gramática, o caráter complexo da nominalização tal como preconizada pelo modelo da GDF. Com um paralelismo notável, a organização morfossintática adotada no formalismo foi capaz de dar uma representação dos papéis que a nominalização exerce, respectivamente, no Nível Interpessoal como estratégia do falante para formular um Subato de Referência ou um Subato de Atribuição, e no Nível Representacional, para designar entidades ou propriedades.

Vimos que Croft (1995) estabelece distinções para o conceito de autonomia, interpretando, por um lado, a autonomia da sintaxe em relação à semântica e à pragmática e, por outro, a autonomia do conhecimento gramatical em relação à orientação social e comunicativa do sistema linguístico. Ao chegar a este balanço, espero ter deixado claro que, embora postulado como um nível autônomo, com suas próprias categorias, o Morfossintático não é inteiramente independente do Interpessoal e do Representacional, justamente porque o conhecimento formal não pode, em minha concepção, ser tomado como independente da orientação social e comunicativa da linguagem.

Se, conforme afirma Du Bois (1985), o grau de preparação de que dispõe o linguista para aferir o significado da teoria se dá a partir de uma perspectiva que vê a gramática como sistema adaptativo, entendo que este trabalho atingiu um nível considerável de significação epistemológica, ou metagramatical. Com efeito, é justamente esse o conceito *lato sensu* de gramática aqui adotado: a relativa autonomia da gramática pressupõe a existência de um sistema e sua relativa suscetibilidade a pressões externas pressupõe postular o caráter adaptativo dos sistemas linguísticos.

Parece-me claro que o tratamento da nominalização permitiu rejeitar concepções autonomistas, cujas limitações na tentativa de explicar a estruturação da linguagem são determinadas apenas pelo reconhecimento de forças motivadoras internas ao sistema. No entanto, foi capaz também de rejeitar concepções funcionalistas

radicais, que são limitadas na tentativa de explicar a preservação de categorias gramaticalizadas em formas mais ou menos congeladas.

Para finalizar, seria bom apelar para a magia da palavra literária, que o Bruxo de Cosme Velho encarna sempre de maneira magistral, como um prestidigitador tão rápido que nem sempre nosso nível de atenção cognitivo logra acompanhar o emaranhado de suas artimanhas.

Na concepção que elabora de Capitu, o Bruxo põe as palavras na boca de um narrador que é também personagem e como tal parte "interessadíssima" no modo como os fatos são narrados. O encantamento dos olhos de ressaca, em que mergulham de corpo e alma marinheiros desavisados como o jovem Bentinho, traz também a natureza oblíqua e dissimulada que o maduro Bentinho passa a enxergar. Se há dois Bentinhos num único – ego e superego ligados pelo id – é uma questão não resolvida, justamente em função do ponto de vista enunciativo que o Bruxo prestidigitador escolhe para nosso deleite.

Nem bruxo, nem prestidigitador, sou apenas aqui o agente responsável pelo fio condutor deste texto, que, à guisa de narrativa, se encerra com um balancete altamente otimista; mas, não custa acrescentar, otimista segundo o ponto de vista do narrador, que, ao fim e ao cabo, é também o autor.

Fica, portanto, para o escrutínio do leitor a tarefa sempre necessária e desejável de uma meticulosa auditoria, que, com base em outra perspectiva menos otimista, pode seguramente descobrir mais perdas do que ganhos. Então o espectro de devedor, no lugar do estatuto de credor como suponho, passa a assombrar este fechamento para balanço, como a desconfiança passou a desconstruir gradativamente o casamento de Bentinho.

A verdade é que a ciência não é uma magia e estar certo ou equivocado é uma questão relativa como, aliás, é verdadeiro também para a categorização linguística. Justamente por ser a verdade uma questão relativa é que "a controvérsia é o sopro da ciência e quando nós todos concordarmos será apenas porque nossa ciência estará morta" (Bach, 1973, p.27).

Referências bibliográficas

ABAURRE, M. B. M.; RODRIGUES, A. C. S. Apresentação. In: ______. (Orgs.). *Gramática do português falado*. Campinas: Editora da Unicamp, 2002. p.9-28.

ANSTEY, M. Functional grammar from its inception. In: MACKENZIE, J. L.; GÓMEZ-GONZÁLEZ, M. Á. (Ed.). *A New Architecture for Functional Grammar*. Berlin: Mouton de Gruyter Amsterdam, 2004. p.23-72.

ASHBY, W. J.; BENTIVOGLIO, P. Information flow in spoken French and Spanish: a comparative study. *Paper* lido no *NWAV 20*, Georgetown University: Washington, 1993.

BACH, E. A linguística estrutural e a filosofia da ciência. In: CHOMSKY, N. et al. *Novas perspectivas linguísticas*. Petrópolis: Vozes, 1973. p.11-27.

BAKER, D. FG expression rules: from templates to constituent structure. *Working Papers in Functional Grammar 67*. Amsterdam: University of Amsterdam, 1999.

______. The FG expression rules: a dynamic model. *Revista Canária de Estudios Ingleses*. La Laguna, v.42, p.15-54, 2001 (Special Issue: Challenges and Developments in Functional Grammar).

BAKER, D.; SIEWIERSKA, A. Towards a speaker model of Functional Grammar. In: MACKENZIE, J. L.; GÓMEZ-GONZÁLEZ, M. Á. (Ed.). *A New Architecture for Functional Grammar*. Berlin: Mouton de Gruyter Amsterdam, 2004. p.325-64.

BASÍLIO, M. *Teoria lexical*. São Paulo: Editora Ática, 1989.

______. Flutuação categorial de base adjetiva no português falado. In: ILARI, R. (Org.). *Gramática do português falado.* Campinas: Editora da Unicamp, 2002. p.75-90 (v.II: Níveis de análise linguística).

______. *Formação e classes de palavras no português do Brasil*. São Paulo: Contexto, 2004.

BENTIVOGLIO, P.; BRAGA, M. L. Espanhol, português e a ordem de palavras. *D.E.L.T.A.* São Paulo. v.4, n.2, p.163-82, 1988.

BENVENISTE, E. *Problemas de linguística geral*. São Paulo: Nacional/ Edusp, 1976.

BOLKESTEIN, M. Limits to layering: locatability and other problems. In: FORTESCUE, M. et al. (Ed.). *Layered Structure and Reference in a Functional Perspective*. Amsterdam: John Benjamins, 1992. p.387-407.

______. What do we do with topic and focus? Evaluating pragmatic information. In: HANNAY, M.; BOLKESTEIN, M. (Ed.). *Functional Grammar and Verbal Interaction*. Amsterdam: John Benjamins, 1998. p.193-214.

BORBA, F. da S. *Uma gramática de valências para o português*. São Paulo: Ática, 1996.

BROWN, R. *Words and Things*. New York: Free Press, 1958.

BUTLER, C. *Structure and Function:* a guide to three major structural-functional theories. Amsterdam: John Benjamins, 2003 (Part I: Approaches to the simplex clause).

BYBEE, J. *Morphology*: a study of the relation between meaning and form. Amsterdam: John Benjamins, 1985.

CAMACHO, R. G.; DALL'AGLIO-HATTNHER, M. M.; GONÇALVES, S. C. L. O substantivo. In: ILARI, R.: NEVES, M. H. M. (Orgs.) *Gramática do português culto falado no Brasil*. Campinas: Editora da Unicamp, 2008. p.21-84 (v.II: Classes de palavras e processos de construção).

CAMACHO, R. G; SANTANA, L. Argument structure of deverbal nouns in Brazilian Portuguese. *Journal of Language and Linguistics.* United Kingdom, v.3, n.2, p.229-42, 2004.

CÂMARA JR., Joaquim M. *Princípios de linguística geral*. 4.ed. ver. e aum. Rio de Janeiro: Livraria Acadêmica, 1972.

CASTELEIRO, J. M. *Sintaxe transformacional do adjetivo*. Lisboa: Instituto Nacional de Investigação Científica, 1981.

CASTILHO, A. T.; PRETI, D. (Org.). *A linguagem falada culta na cidade de São Paulo*: materiais para seu estudo. São Paulo: T. A. Queiroz, 1986. v.1.

______. (Org.). *A linguagem falada culta na cidade de São Paulo*: materiais para seu estudo. São Paulo: T. A. Queiroz, 1987. v.2.

CASTILHO, A. T. *A língua falada no ensino do português*. São Paulo: Contexto, 1998.

______. *Proposta funcionalista de mudança linguística*. Os processos de lexicalização, semanticização, discursivização e gramaticalização na constituição das línguas. Campinas, 2003 (inédito).

______. Apresentação. In: JUBRAN, C. C. A. S.; KOCH, I. G. V. (Orgs.). *Gramática do português falado culto no Brasil*. Campinas: Editora da Unicamp, 2006. (v.I: Construção do texto falado).

CHAMBERS, J. K. *Sociolinguistic Theory*: linguistic variation and its social significance. Oxford: Blackwell, 1996.

CHOMSKY, N. Remarks on nominalization. In: JACOBS, R. A., ROSENBAUM, P. (Ed.). *Readings in the Theory of Transformational Grammar*. Waltham: Ginn, 1970.

______. *Aspectos da teoria da sintaxe*. Trad. de J. A. Meireles e E. Raposo. Coimbra: Armênio-Amado Editor, 1975.

______. *Rules and Representation*. New York: Columbia University Press, 1980.

______. *Language Universals and Linguistic Typology*. Oxford: Blackwell, 1989.

COMRIE, B; THOMPSON, S. Lexical nominalizations. In: SHOPEN, T. (Ed.). *Language Typology and Syntactic Description*. Cambridge: Cambridge University Press, 1985. p.349-98. v.3.

CONNOLY, J. The question of discourse representation in Functional Discourse Grammar. In: MACKENZIE, J. L.; GÓMEZ-GONZÁLEZ, M. Á. (Ed.). *A New Architecture for Functional Grammar*. Berlin: Mouton de Gruyter Amsterdam, 2004. p.89-116.

CRISTOFARO, S. *Subordination*. Oxford: Oxford University Press, 2003.

CROFT, W. *Syntactic Categories and Grammatical Relations:* the cognitive organization of information. Chicago: University of Chicago Press, 1991.

______. Autonomy and functionalist linguistics. *Language*, Baltimore. v.71, p.490-532, 1995.

DIK, S. C. Some basic principles of Functional Grammar. In: HATTORI, S. et al. (Ed.). *Proceedings of the XIIIth International Congress of Linguists*. Tokyo: Tokyo Press, 1983. p.74-88.

______. Formal and semantic adjustment of derived constructions. In: BOLKSTEIN et al. (Ed.). *Predicates and Terms in Functional Grammar*. Dordrecht: Foris, 1985. p.1-28.

______. *The Theory of Functional Grammar*. Dordrecht: Foris, 1989 (Part I: The structure of the clause).

______. *The Theory of Functional Grammar*. Edited by Kees Hengeveld. Berlin: Mouton de Gruyter, 1997 (Part II: Complex and Derived Constructions).

DIXON, R . L. W. Ergativity. *Language*, Baltimore, v.55, p.59-138, 1979.

DU BOIS, J. W. Beyond definitess: the trace of identity in discourse. In: CHAFE, W. (Ed.). *The Pear Stories*. Norwood, NJ: Ablex, 1980. p.203-74.

______. Competing motivations. In: HAIMAN, J. (Ed.). *Iconicity in Syntax*. Amsterdam: John Benjamins, 1985. p.343-65.

______. The discourse basis of ergativity. *Language,* Baltimore, v.63, p.805-55, 1987.

DUTRA, R. The hybrid S-category in Brazilian Portuguese: some implications for word order. *Studies in Language*. Philadelphia, v.11, n.1, p.163-80, 1987.

FOLEY, W. A.; VAN VALIN, R. D. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press, 1984.

GARCIA, E. Discourse without syntax. In: GIVÓN, T. (Ed.). *Syntax and Semantics*. New York: Academic Press, 1979 (v.12: Discourse and syntax).

GARCIA VELASCO, D. *Funcionalismo y linguística:* la Gramática Funcional de S. C. Dik. Oviedo, Universidad de Oviedo, 2003.

______; HENGEVELD, K. Do we need predicate frames? In: MAIRAL USÓN, R.; QUINTERO, M. J. P. (Ed.). *New Perspectives on Argument Structure in Functional Grammar,* Berlin: Mouton de Gruyter, 2002. p.95-123.

______. PORTERO MUÑOZ, C. Understood objects in Functional Grammar. *Working Papers in Functional Grammar*, v.76. Amsterdam: University of Amsterdam, 2002.

GIVÓN, T. *On Understanding Grammar*. New York: Academic Press, 1979.

______. The binding hierarchy and the typology of complements. *Studies in Language*, v.4, n.3, p.333-77, 1980.

______. Logic vs. pragmatics with natural language as the referee. *Journal of pragmatics*, v.6, n.2, p.81-133, 1981.

______. *Syntax:* a functional-typological introduction. Amsterdam: John Benjamins, 1984. (v.1)

______. *Syntax:* a functional-typological introduction. Amsterdam: John Benjamins, 1990. (v.2)

GRICE, H. P. Logic and conversation. In: COLE, P., MORGAN, J. L. (Ed.). *Syntax and Semantics.* New York: Academic Press, 1975. p.41-58 (v.3: Speech Acts).

HAIMAN, J. Iconic and economic motivation. *Language*, Baltimore, v.59, p.781-819, 1983.

______. The iconicity of grammar. *Language*, Baltimore, v.56, p.515-40, 1987.

HALLIDAY, M.A.K. Language structure and language functions. In: LYONS, J. (Ed.). *New Horizons in Linguistics*. Harmondsworth: Penguin, 1970. p.140-65.

______. *An Introduction to Functional Grammar*. London, Edward Arnold, 1985.

HANNAY, M.; BOLKESTEIN, M. Introduction. In: HANNAY, M.; BOLKESTEIN, M. (Ed.). *Functional Grammar and Verbal Interaction*. Amsterdam: John Benjamins, 1997. p.vii-xii.

HARDER, P. Comment clauses, Functional Discourse Grammar and the grammar-discourse interface. In: MACKENZIE, J. L.; GÓMEZ-GONZÁLEZ, M. Á. (Ed.). *A New Architecture for Functional Grammar.* Berlin: Mouton de Gruyter Amsterdam, 2004, p.197-210.

______. *Functional Semantics.* a theory of meaning, structure and tense in English. Berlin: Mouton de Gruyter, 1996.

HASPELMATH, M. Word-class-changing inflection and morphological theory. *Yearboof of Morphology*. Kluwer Academic Publishers, 1995, p.43-66.

HEINE, B.; CLAUDI, U; HÜNNEMEYER, F. *Grammaticalization*: a conceptual framework. Chicago: The University of Chicago Press, 1991.

HENGEVELD, K. Clause structure and modality in Functional Grammar. In: AUWERA, J. Van der; GOSSENS, L. (Ed.). *Ins and Outs of Predication*. Dordrecht: Foris, 1987. p.53-66.

______. Illocution, mood and modality in a functional grammar of Spanish. *Journal of Semantics*, v.6, p.227-69, 1988.

______. Layers and operators in Functional Grammar. *Journal of Linguistics*, v.25, p.125-57, 1989.

______. *Non-verbal predication*: theory, typology, diachrony. Berlin: Mouton de Gruyter, 1992.

______. Cohesion in Functional Grammar. In: CONNOLY et al. (Ed.). *Discourse and Pragmatics in Functional Grammar*. Berlin and New York: Mouton de Gruyter, 1997. p.1-16.

______. Formalizing functionally. In: DARNELL, M. et al. (Ed.). *Functionalism and Formalism in Linguistics*. Amsterdam: John Benjamins, 1998. p.93-105, v.2: (Studies in Language Companion Series 42).

______. The architecture of a Functional Discourse Grammar. In: MACKENZIE, J. L.; GÓMEZ-GONZÁLEZ, M. Á. (Ed.). *A New Architecture for Functional Grammar*. Berlin: Mouton de Gruyter, 2004a. p.1-21 (Functional Grammar Series 24).

______. Epilogue. In: MACKENZIE, J. L.; GÓMEZ-GONZÁLEZ, M. Á. (Ed.). *A New Architecture for Functional Grammar*. Berlin: Mouton de Gruyter, 2004b. p.365-78 (Functional Grammar Series 24).

______. Functional Discourse Grammar. In *Functional/Generative Grammar*. Curso de Pós-Graduação. Universidade de Amsterdã, 2004c.

______. State-of-Affairs concepts. In: BOOIJ, G.; LEHMANN, C.; MUGDAN, J. (Ed.). *Morphology:* a handbook on inflection and word formation. Berlin: Mouton de Gruyter, 2004d. v.2, p.1104-111.

______. Dynamic Expression in Functional Grammar. In: DE GROOT, Casper; HENGEVELD, Kees (Ed.). *Morphosyntactic Expression in Functional Grammar*. Berlin: Mouton de Gruyter, 2005. p.53-86.

______; MACKENZIE, J. L. *Functional Discourse Grammar*: a typologically-based theory of language structure. Oxford: Oxford University Press, 2008.

______. Functional discourse grammar. In: BROWN, K. (Ed.). *Encyclopedia of Language and Linguistics*. Oxford: Elsevier, 2006. p.668-76.

HJELMLSLEV, L. *Prolegômenos a uma teoria da linguagem*. São Paulo: Perspectiva, 1973.

HOEKSTRA, T. Deverbalization and inheritance. *Linguistics*, v.24, n.3, p.549-84, 1986.

HOPPER, P. J. The iconicity of the universal categories "noun"and "verb". In: HAIMAIN, J. (Ed.). *Iconicity in Syntax* (Proceedings of a Symposium on Iconicity in Syntax). Amsterdam/Philadelphia: John Benjamins, 1985.

______. Emergent grammar. *Berkeley Linguistics Society,* v.13, p.139-57, 1987.

HOPPER, P. J.; THOMPSON, S. A. The discourse basis for lexical categories in universal grammar. *Language*. Baltimore, v.60, p.703-52, 1984.

HOPPER, P. J.; TRAUGOTT, E. C. *Grammaticalization*. Cambridge: Cambridge University Press, 1993.

HYMES, D. On communicative competence. In PRIDE, J. B.; HOLMES, J. (Ed.). *Sociolinguistics*. Harmondsworth: Penguin, 1972. p.269-93.

HYMES, D. Short notice of Dik's Studies in Functional Grammar. *Language in Society*, v.12, 1979.

ILARI, R.; NEVES, M. H. M. (Orgs.). *Gramática do português culto falado no Brasil*. Campinas: Editora da Unicamp, 2008. (v.II: Classes de palavras e processos de construção).

ILARI, R. et al. Considerações sobre a posição dos advérbios. In: CASTILHO, A. T. (Org.). *Gramática do português falado*. 4.ed. ver. Campinas: Editora da Unicamp/Fapesp, 2002. p.53-120. v.I: A ordem.

JUBRAN, C. C. A. S.; KOCH, I. G. V. *Gramática do português falado culto no Brasil*. Campinas: Editora da Unicamp, 2006. (v.I: Construção do texto falado).

KEIZER, E. English Prepositions in FGD. Paper presented at *11th International Conference on Functional Grammar*. Gijón, Espanha, 2004a.

______. Term structure in FG: a modest proposal. *Working Papers in Functional Grammar.* V. 78. University of Amsterdam, 2004b.

KOPTJEVSKAJA-TAMM, M. *Nominalizations*. London: Routledge, 1993.

KROON, C. Discourse markers, discourse structure and Functional Grammar. In: CONNOLLY, J. et. al. (Ed.). *Discourse and Pragmatics in Functional Grammar*. Berlin: Mouton de Gruyter, 1997. p.17-32.

JACKENDOFF, R. *Semantic Structures*. Cambridge: MIT Press, 1990.

LANGACKER, R. W. Nouns and verbs. *Language*. Baltimore, v.63, 1987.

LEHMANN, C. Towards a typology of clause linkage. In: HAIMAN, J.; THOMPSON, S. A. (Ed.). *Clause Combining in Grammar and Discourse*. Amsterdam: John Benjamins, 1988.

LIEDTKE, F. Illocution and grammar: a double level approach. In: HANNAY, M.; BOLKESTEIN, M. (Ed.). *Functional Grammar and Verbal Interaction*. Amsterdam: John Benjamins, 1998. p.107-27.

LYONS, J. *Semantics*. Cambridge: Cambridge University Press, 1977.

______. *Introdução à linguística teórica*. São Paulo: Ed. Nacional, 1979.

______. *Lingua(gem) e linguística:* uma introdução. Rio de Janeiro: Zahar, 1982.

MACKENZIE, J. L. Nominalization and valency reduction. In: BOLKESTEIN, A. M. et al. (Ed.). *Predicates and Terms in Functional Grammar*. Dordrecht: Foris Publications, 1985. p.31-51.

______. Nominalizations and basic constituent ordering. In: VAN DER AUWERA, J.; GOOSSENS, L. (Ed.). *Ins and Outs of the Predication*. Dordrecht: Foris Publications, 1987. p.93-106.

______. English nominalizations in the layered model of the sentence. In: DEVRIEND, B.; GOOSSENS, L.; AUWERA, J. van der (Ed.). *Complex Structures:* a functionalist perspective. Berlin: Mouton de Gruyter, 1996. p.325-55.

______. The basis of syntax in the holophrase. In: HANNAY, M.; BOLKESTEIN, M. (Ed.). *Functional Grammar and Verbal Interaction*. Amsterdam: John Benjamins, 1998. p.267-96.

______. FGD perspectives on nominalization. Paper presented at the *11th International Conference on Functional Grammar*. Gijón, Espanha, 2004a.

______. Functional Discourse Grammar and language production. In: MACKENZIE, J. L.; GÓMEZ-GONZÁLEZ, M. Á. (Ed.). *A New Architecture for Functional Grammar*. Berlin: Mouton de Gruyter Amsterdam, 2004b. p.179-96 (Functional Grammar Series 24).

MAINGUENEAU, D. L'unité de la linguistique. *D.E.L.T.A*. São Paulo, v.6, n.2, p.127-38, 1990.

MALCHUKOV, A. L. *Nominalization/Verbalization:* constraining a typology of transcategorial operations. Lincom: Lincom Europa, 2004.

MANN, W. C.; THOMPSON, S. A. Rhetorical structure theory: a framework for the analysis of texts. *Papers in Pragmatics*, v.1, p.79-105, 1987.

MARTINET, A. *Elementos de linguística geral.* Lisboa: Sá da Costa, 1972.

MÜLLER, A. L. et al. Adjetivos no português do Brasil: predicados, argumentos ou quantificadores? In: ABAURRE, M. B .M; RODRIGUES, A. C. (Org.). *Gramática do português falado*. Campinas: Editora da Unicamp/Fapesp, 2002. p.317-44 (v.VIII: Novos estudos descritivos).

NEVES, M. H. M. O substantivo (ou nome) comum. In: VII SEMINÁRIO DO PROJETO GRAMÁTICA DO PORTUGUÊS FALADO, 1993, Campos do Jordão.

______. Estudo das construções com verbo-suporte em português. In: KOCH, I. G. V. (Org.). *Gramática do português falado*. Campinas: Editora da Unicamp, 2002. p.209-38. (v.VI: Desenvolvimentos)

______. Estudo da estrutura argumental dos nomes. In: KATO, M. (Org.). *Gramática do português falado*. Campinas: Editora da Unicamp/São Paulo: Fapesp, 1996. p.119-54 (v.5: Convergências).

NEWMEYER, E. F. Some remarks on the Functionalist-Formalist controversy in linguistics. In: DARNELL, M. et al. (Ed.). *Functionalism and Formalism in Linguistics.* Amsterdam: John Benjamins, 1998. p.469-86 (v.2: Studies in Language Companion Series 42).

NOONAN, M. Complementation. In: SHOPEN, T. (Ed.). *Language Typology and Syntactic Description*. Cambridge: Cambridge University Press, 1985. p.42-141 (v.2: Complex constructions).

______. Non-structuralist syntax. In: DARNELL, M. et al. (Ed.). *Functionalism and Formalism in Linguistics.* Amsterdam: John Benjamins, 1998. p.11-31 (v.2: Studies in Language Companion Series 42).

NUYTS, J. *Aspects of a Cognitive-Pragmatic Theory of Language*. Amsterdam: John Benjamins, 1992.

PEZATTI, E. G. *A ordem de palavras em português: aspectos tipológicos e funcionais*. Araraquara, 1992. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista.

PRETI, D.; URBANO, H. (Org.). *A linguagem falada culta na cidade de São Paulo*. São Paulo: T. A. Queiroz/Fapesp, 1988. (v.3: Materiais para seu estudo).

PRINCE, H. F. Toward a taxonomy of given-new information. In: COLE, P. (Ed.). *Radical Pragmatics.* New York: Academic Press, 1981. p.223-55.

RAMIREZ, H. *A fala tukano dos ye'pâ-masa:* gramática, dicionário, método de aprendizagem. Manaus: Inspetoria Salesiana da Amazônia, Cedem, 1997.

RAPOSO, E. P. *Teoria da gramática:* a faculdade da linguagem. Lisboa: Editora Caminho, 1992.

RIJKHOFF, J. *The Noun Phrase*. Oxford: Oxford Univesity Press, 2002.

ROSCH, E. H. Natural categories. *Cognitive Psychology.* v.4, n.3, p.328-50, May 1973.

SANTANA, L. *A expressão da estrutura argumental dos nomes derivados*. São José do Rio Preto, 2005. Dissertação (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista.

SAPIR. E. *A linguagem:* introdução ao estudo da fala. Rio de Janeiro: Livraria Acadêmica, 1971.

SAUSSURE, F. de. *Curso de linguística geral*. São Paulo: Cultrix, 1977.

SILVA, A. S. da. A linguística cognitiva: uma breve introdução a um novo paradigma em linguística. *Revista Portuguesa de Humanidades.* Braga, v.1, n.1-2, p.59-101, 1997.

SINCLAIR, J. A., COULTHARD, R. M., *Towards an Analysis of Discourse:* the English used by teachers and pupils. London: Oxford University Press, 1975.

STEUTEN, A. A. G. Structure and coherence in business conversations. In: HANNAY, M.; BOLKESTEIN, M. (Ed.). *Functional Grammar and Verbal Interaction*. Amsterdam: John Benjamins, 1998. p.59-75.

TAYLOR, J. R. *Linguistic Categorization*: prototypes in linguistic theory. Oxford: Oxford University Press, 1989.

TRAUGOTT, E. C.; HEINE, B. (Org.). *Approaches to Grammaticalization*. Amsterdam: John Benjamins, 1991.

VET, C. The multilayered structure of the utterance: about illocution, modality and discourse moves. In: HANNAY, M.; BOLKESTEIN, M. (Ed.). *Functional Grammar and Verbal Interaction*. Amsterdam and Philadelphia: John Benjamins, 1998. p.1-23.

WEEDWOOD, B. *História concisa da linguística*. São Paulo: Parábola, 2003.

SOBRE O LIVRO

Formato: 14 x 21 cm
Mancha: 23,7 x 42,5 paicas
Tipologia: Horley Old Style 10,5/14
Papel: Off-set 75 g/m² (miolo)
Cartão Supremo 250 g/m² (capa)
1ª edição: 2011

EQUIPE DE REALIZAÇÃO

Coordenação Geral
Marcos Keith Takahashi

Neste livro, Roberto Gomes Camacho enfrenta um problema clássico do Funcionalismo: a questão da categorização linguística, focalizada no modelo em sua natureza contínua, não opositiva, não negativa. Surgiu daqui a teoria dos protótipos, já bem conhecida no Brasil. Considerando o lado complexo, no sentido de "interligado", o autor toma como parâmetro as diferenças entre o nome e o verbo, situando a meio caminho o importante processo da nominalização, "postulada como constituindo um termo complexo entre os dois outros opostos, tomados como membros prototípicos". Para comprovar essa postulação, o autor argumenta que as expressões nominalizadas partilham propriedades nominais e verbais, comprovadas, neste caso, pela manutenção da estrutura argumental encontrada no verbo.

Com este trabalho, Roberto Gomes Camacho atinge sua plena maturidade intelectual, duramente conquistada ao longo de uma carreira acadêmica exemplar.

ISBN 978-85-393-0189-8